

Leticia Bolian Zimback



Mestranda em Ecologia, LabTrop, Instituto de Biociências, USP

Meus Exercícios

[Exercícios Leticia Bolian Zimback](#)

Proposta de Trabalho Final

Plano A

Meu projeto de mestrado consiste em avaliar a chuva de sementes em uma área de Restinga. Portanto, pensei em criar uma função que fizesse uma análise exploratória e me desse algumas informações iniciais sobre os meus dados.

Esta função teria apenas dois argumentos. O primeiro argumento seria um `data.frame` com as variáveis categóricas como, por exemplo, espécie, mês, ano, parcela; e a variável resposta que é a quantidade de sementes. O segundo argumento teria a função de retirar os dados faltantes do meu `data.frame`. Dentro da função quero calcular várias coisas como: a quantidade de sementes em cada mês, a quantidade de sementes em cada parcela, a quantidade de sementes em cada mês/parcela, o número médio de sementes no ano, número médio de sementes no mês, gráficos que testem a homogeneidade na variância dos meus dados e a normalidade dos resíduos, outros gráficos de análise exploratória como `boxplot` e `histograma`, análise da variância do meu conjunto de dados e, por fim, testar modelos para ver quais das variáveis categóricas explicam melhor a quantidade de sementes que cai nos coletores.

Plano B

Além do mestrado, eu trabalho na Prefeitura Municipal de São Paulo com levantamento e monitoramento de fauna na cidade. Nós recebemos animais silvestres machucados e também de apreensão.

Pensei em fazer uma função que quantifique os animais recebidos provenientes de cada uma das regiões da cidade (zonas norte, sul, leste, oeste, centro), divididos pelo fator classe (aves, mamíferos, répteis e anfíbios). Além disto quero plotar estes pontos em um mapa da cidade, junto com o valor médio de animais recebidos por classe em cada uma das regiões. Quero fazer isto separadamente para os animais que são recebidos machucados e depois para os apreendidos.

Comentários

Comentários Leticia, acho que a primeira proposta está mais trabalhada, a segunda foi meio que para constar... vc. não explica direito o que quer. Portanto, não pude contar com esse plano B! O plano A precisa de alguns acabamentos. A primeira parte é uma manipulação do data-frame para conseguir os dados síntese das espécies por mês (ótimo), a segunda parte está um pouco frouxa. Defina melhor uma análise exploratória para esse tipo de dado e fique só com isso! Minha sugestão é que esqueças os gráficos diagnósticos de normalidade e homogeneidade de variância, fazem mais sentido quando já se tem um modelo em mente, o que não é o caso. Normalmente dados fenológicos são analisados com *estatística circular* e são representado de maneira muito diferente da usual em modelos lineares. Sugiro que dê uma olhada nos pacotes “circular” e “plotrix” e inclua na sua função além da parte da tratar os dados brutos faça gráficos síntese da fenologia da área por mês na forma de estatística circular. Veja um exemplo:



Que está em um artigo bem simples na: [Brazilian Journal of Biology](#)

Aguardamos suas correções! Avise-me quando terminar pelo email abaixo. —
[Alexandre Adalardo de Oliveira](#) 2014/04/25 17:02

Proposta A

Pensando nas dicas propostas, resolvi simplificar um pouco as análises iniciais para dar maior ênfase para o gráfico de estatística circular.

A função terá como input dois argumentos. O primeiro será um data.frame com quatro colunas, que devem ser o mês, o ano, a parcela coletada e a quantidade de sementes (variável resposta). O segundo argumento irá retirar os dados faltantes do data.frame. Através da indexação das colunas correspondentes a função irá calcular o número de sementes por mês, o número de sementes por parcela, o número médio de sementes no ano e o número médio de sementes no mês.

A partir deste número médio de sementes no mês, será utilizada uma estatística circular para plotar os dados e verificar se há variação anual na chegada de sementes nesta comunidade vegetal.

Página de Ajuda

pheno.sementes

package:nenhum

R Documentation

Análise da fenologia da chuva de sementes

Description:

Através de um `data.frame` padronizado calcula alguns dados brutos e gera um gráfico circular através do uso de estatística circular.

Usage:

```
feno.sementes(x, rmNA=TRUE)
```

Arguments:

`x`: Objeto da classe 'data.frame' padronizado. Para mais detalhes (ver 'Details').

`rmNA`: lógico. Remover dados faltantes do `data.frame`.

Details:

O `data.frame` (`x`) deve ser composto de quatro colunas. A primeira coluna diz respeito aos meses do estudo, a segunda aos anos, a terceira às parcelas amostradas e a quarta é a variável resposta que corresponde a quantidade de sementes encontradas no estudo.

A partir do `data.frame` são calculados alguns dados brutos, estes são: o total de sementes que chegou ao longo do estudo, soma das sementes que chegaram em cada mês, soma das sementes que chegaram em cada parcela, soma das sementes que chegaram em cada ano, média das sementes que chegaram por ano e média das sementes que chegaram em cada um dos meses.

Através do pacote "circular" são calculados alguns parâmetros de estatística circular.

O objeto calculado que contém a média das sementes que chegaram em cada um dos meses é utilizado e transformado em dados circulares. A partir desta transformação são obtidos a direção média dos dados e o comprimento deste vetor médio. Além desta

informação, dois testes (Rayleigh.teste e Watson.teste) são feitos para verificar a uniformidade dos dados.

Value:

Um gráfico é gerado. Neste gráfico os eixos são os meses e estão dispostos na circunferência, o vetor médio é plotado de acordo com o seu comprimento e direção média dados.

Além do gráfico a função retorna todos os cálculos brutos e os valores da média circular, comprimento deste vetor média e o resultado dos testes.

Author(s):

Leticia Bolian Zimback

leticia.zimback@gmail.com

References:

Garcia, L.C., Barros, F.V., Lemos-Filho, J.P. (2009) Fructification phenology as an important tool in the recovery of iron mining areas in Minas Gerais, Brazil. *Brazilian Journal of Biology* 69(3): 887-893.

Lund, U., Agostinelli, C. (2013). Package 'circular'. Disponível online em <http://cran.r-project.org/web/packages/circular/circular.pdf>

Ribas, A. (2013). *Recologia – Errando em Ecologia e Evolução – Gráficos circulares*. Disponível online em <http://recologia.com.br/2013/04/graficos-circulares/>

Sfair, J.C. Estatística Circular – Principais análises utilizando o R.

Disponível online em

http://www.researchgate.net/publication/233924997_Estatstica_Circular_-_Principais_anlises_utilizando_o_R

See Also:

Para outra forma de fazer um gráfico circular, veja a função 'radial.plot' do pacote plotrix.

Examples:

#Como a função se comporta quando há dados faltantes (NAs):

```
mes <-
c("jan","jan","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","m
ai","mai","mai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set",
"set","set","out","out","out","nov","nov","nov","dez","dez","dez","jan","jan
","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","mai","mai","m
ai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set","set","set",
"out","out","out","nov","nov","nov","dez","dez","dez")
```

```
ano <-
c("2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","201
2","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","20
12","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","2
012","2012","2012","2012","2013","2013","2013","2013","2013","2013","2013","
2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013",
"2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013",
"2013","2013","2013","2013","2013","2013")
```

```
parcela <-
c("1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1",
"2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2",
"3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3",
"1","2","3","1","2","3","1","2","3","1","2","3","1","2","3")
```

```
sementes <-
c(4,NA,7,4,4,9,2,10,4,7,3,8,8,2,NA,2,6,1,9,5,9,NA,2,11,3,4,9,6,3,7,1,NA,2,7,
5,2,8,1,NA,3,2,7,8,9,2,4,4,7,3,8,2,4,1,5,1,8,6,NA,NA,2,8,9,10,6,2,3,5,5,6,3,
```



```

mes <-
c("jan","jan","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","m
ai","mai","mai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set",
"set","set","out","out","out","nov","nov","nov","dez","dez","dez","jan","jan
","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","mai","mai","m
ai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set","set","set",
"out","out","out","nov","nov","nov","dez","dez","dez")

ano <-
c("2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","201
2","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","20
12","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","2
012","2012","2012","2012","2013","2013","2013","2013","2013","2013","2013","
2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013",
"2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013
","2013","2013","2013","2013","2013","2013","2013")

parcela <-
c("1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1
","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2
","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3
","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3")

sementes3 <-
c(45,76,87,98,76,87,65,54,65,67,76,87,54,43,56,87,76,89,95,75,91,66,81,99,45
,43,56,54,43,45,65,76,78,89,87,76,46,77,88,99,77,88,66,55,66,87,43,59,78,76,
55,69,78,91,80,83,60,78,93,90,45,44,66,55,49,56,66,77,79,90,88,77)

z <- data.frame(mes,ano,parcela,sementes3)

feno.sementes(z,rmNA=TRUE)

#Quando a abundância em alguns meses é igual a zero:

mes <-
c("jan","jan","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","m
ai","mai","mai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set",
"set","set","out","out","out","nov","nov","nov","dez","dez","dez","jan","jan
","jan","fev","fev","fev","mar","mar","mar","abr","abr","abr","mai","mai","m
ai","jun","jun","jun","jul","jul","jul","ago","ago","ago","set","set","set",
"out","out","out","nov","nov","nov","dez","dez","dez")

ano <-
c("2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","201
2","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","20
12","2012","2012","2012","2012","2012","2012","2012","2012","2012","2012","2
012","2012","2012","2012","2013","2013","2013","2013","2013","2013","2013","
2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013",
"2013","2013","2013","2013","2013","2013","2013","2013","2013","2013","2013
","2013","2013","2013","2013","2013","2013","2013")

```

```
parcela <-  
c("1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3","1",  
"2","3","1","2","3","1","2","3","1","2","3","1","2","3","1","2",  
"3","1","2","3","1","2","3","1","2","3","1","2","3","1","2","3",  
"1","2","3","1","2","3","1","2","3","1","2","3","1","2","3")  
  
sementes4 <-  
c(0,0,0,0,0,0,0,0,0,10,43,32,54,21,43,12,16,29,43,49,29,48,23,43,0,0,0,0,0,0,  
0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,12,45,34,56,23,45,14,18,31,45,51,31,50,25,45,  
0,0,0,0,0,0,0,0,0,0,0,0,0)  
  
a <- data.frame(mes,ano,parcela,sementes4)  
  
feno.sementes(a,rmNA=TRUE)
```

Código da Função

```
feno.sementes<-function(x,rmNA=TRUE)  
{  
  
##Coloquei um aviso inicial para que as pessoas que utilizarem a função  
tenham o cuidado de importar o data.frame de forma que este se adeque a  
função:  
  
warning("Para que a função rode de forma correta o data.frame deve ter  
quatro colunas.\nA primeira são os meses do ano. \nA segunda são os anos de  
estudo. \nA terceira são as parcelas. \nA quarta é quantidade de sementes  
que é a variável resposta.\n", call.=FALSE,immediate.=TRUE)  
  
##Comandos para que os dados faltantes sejam excluídos da análise e isso  
seja explicitado no Console:  
  
if(rmNA==TRUE)  
  
  {  
  
    dados=(na.omit(x))  
  
#Como é um data.frame calculei a diferença no número de linhas, pois o  
length é igual para todas, é o número de colunas:  
    dim <- dim(x)-dim(dados)  
  
    n.NA <- dim[1]  
    cat("Valores NA excluídos\n",n.NA,"\n","\n")  
  
  }  
  
}
```



```
}  
  
else  
  
{  
  
  dados=x  
  
  warning("Se o seu data.frame tiver valores faltantes, a função  
  não irá rodar. O argumento rmNA deve ser verdadeiro\n",  
  call.=FALSE,immediate.=TRUE)  
  }  
  
##Ordenei os níveis do fator mês:  
  
dados[,1] <-  
factor(dados[,1],levels=c("jan","fev","mar","abr","mai","jun","jul","ago","s  
et","out","nov","dez"))  
  
##Transformei as variáveis ano e parcela em fatores:  
  
dados[,2] <- factor(dados[,2])  
  
dados[,3] <- factor(dados[,3])  
  
##Criei este objeto apenas para ajudar na hora de plotar o gráfico circular:  
  
meses <-  
c("jan","fev","mar","abr","mai","jun","jul","ago","set","out","nov","dez")  
  
##Alguns cálculos brutos:  
  
##Cálculo do total de sementes que chegaram ao longo do estudo:  
  
soma.total <- sum(dados[,4])  
  
##Cálculo de quantas sementes chegaram em cada mês:  
  
soma.mes <- tapply(dados[,4],dados[,1],sum)  
  
##Cálculo de quantas sementes chegaram em cada parcela:  
  
soma.parcela <- tapply(dados[,4],dados[,3],sum)  
  
##Cálculo de quantas sementes chegaram em cada ano:  
  
soma.ano <- tapply(dados[,4],dados[,2],sum)  
  
##Cálculo da média das sementes que chegam por ano no estudo:  
  
media.ano <- round(mean(soma.ano),0)
```

```
##Cálculo da média de sementes que chegaram em cada um dos meses:

feno <- round(tapply(dados[,4],dados[,1],mean),0)

##Instalação e abertura do pacote circular. Utilizei a função 'is.element'
para conseguir obter um vetor com o nome dos pacotes instalados. Criei a
função 'instalados' para verificar se o pacote 'circular' já está instalado
ou não:
  instalados <- function(pacotes)
  {
    is.element(pacotes, installed.packages()[,1])
  }
if(instalados("circular")==FALSE)
  {
    install.packages("circular")
  }
if(instalados("circular")==TRUE)
  {
    require("circular", quietly=TRUE)
  }
##Fiz um vetor de doze números espaçados igualmente que serão posicionados
radialmente na circunferência:

angulo <- seq(0,330,by=30)

##Transformei este vetor em valores radianos:

angulo.rad <- rad(angulo)

##Transformei os dados referentes a média de sementes que chegaram por mês
(feno) em valores circulares, relacionados com o vetor angulo.rad criado
anteriormente. Coloquei no sentido anti-horário, pois não consegui mudar
este padrão nas funções posteriores:

feno.c <-
circular(rep(angulo.rad,feno),units=c("radians"),rotation="counter")

##Calculei a média para obter o vetor que indica a direção das observações.
Este valor vai de 0 (início na posição 3h) até 2pi(~6.283185):

media <- mean(feno.c)

##Calculei o comprimento do vetor médio. Quanto mais perto de zero, mais
abrangente o período fenológico ao longo do ano, os dados não tem um padrão
muito evidente:

r <- rho.circular(feno.c)

##Testes de Uniformidade.
```

```
##Valores significativos do p-valor indicam a existência de um padrão e
menor abrangência fenológica:

Rayleigh.teste <- rayleigh.test(feno.c)

##Valores acima do valor crítico, indicam a existência de um padrão e menor
abrangência fenológica:

Watson.teste <- watson.test(feno.c,alpha=0.05)

##Plotei os dados em um gráfico circular:

plot(feno.c,axes=F,main="Fenologia da Chuva de Sementes")

##Coloquei os meses no eixo:

axis.circular(at=circular(rad(angulo)),labels=meses,units=c("radians"))

##Plotei este diagrama para visualizar como a chegada de sementes se
distribui ao longo do ano:

rose.diag(feno.c,bins=12,add=T,axes=F)

##Esta função plota um vetor na direção média dos dados e em um comprimento
que mostra o quanto esta direção influencia os dados como um todo(r):

arrows.circular(x=media, shrink=r, col="red",length=0.1)

##Fiz uma lista para que os resultados sejam apresentados no Console:

lista <-
list(soma.total,soma.mes,soma.parcela,soma.ano,media.ano,feno,media,r,Raylei
gh.teste,Watson.teste)

names(lista) <- c("Total de sementes que chegaram ao longo do estudo","Soma
das sementes que chegaram em cada mês","Soma das sementes que chegaram em
cada parcela","Soma das sementes que chegaram em cada ano","Média das
sementes que chegaram por ano no estudo","Média de sementes que chegaram em
cada um dos meses","Direção média dos dados","Comprimento médio dos
dados","Teste de Uniformidade de Rayleigh","Teste de Uniformidade de
Watson")

return(lista)

}
```

Arquivo da Função

[fenosementesfinal.r](#)

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

Permanent link: 
http://labtrop.ib.usp.br/doku.php?id=cursos:ecor:05_curso_antigo:r2015:alunos:trabalho_final:leticia.zimback:start

Last update: **2020/07/27 18:48**