

Alex Hubbe

Alex Hubbe

Doutorando do Departamento de Genética e Biologia Evolutiva - IB/USP [alex_ex_7_f.r](#)

exercícios

[alex_ex1_f.r](#) [alex_ex2_f.r.txt](#) [alex_ex3_f.r](#) [alex_ex4_f.r](#) [alex_ex5_f.r](#) [alex_ex6_f.r](#) [alex_ex_7_f.r](#)
[alex_ex8_f.r](#) [alex_ex9_f.r](#)

Proposta de trabalho final

Plano A - Meu doutorado está baseado na coleta de pontos em três dimensões de crânios de mamíferos. Para verificar o erro associado a esta coleta cada crânio é medido duas vezes. Após a dupla medição, uma abordagem inicial para avaliar este erro é visualizar graficamente a distância entre os pontos homólogos para as duas medições. O objetivo do plano A é realizar uma função que faça uma triagem dos espécimes medidos que apresentem uma diferença entre os pontos homólogos nas duas medições maior que algum valor a ser definido. Por exemplo, dos x crânios amostrados, a função indicará quais deles apresentam uma ou mais variáveis (pontos) com diferença entre as duas medidas maior que y . Assim, não precisarei olhar crânio por crânio a procura de grandes diferenças! Irei direto ao ponto!

Plano B - um plano nada biológico

Desenvolver uma função que calcule, dada a velocidade média para as diferentes modalidades do triathlon (natação, ciclismo e corrida) e as distâncias a serem percorridas em cada modalidade, o tempo total de prova e os tempos parciais (para cada modalidade).

Comentários PI

Ótimo plano A. Dá para ficar com ele, que é bem mais interessante e útil para vc que o B.

Tópicos específicos

- Pontos homólogos ou os mesmos pontos em cada crânio?
- Além de deixar o usuário escolher o limite de tolerância, vc pode deixar algumas rotinas automáticas, como o critério de identificação de outliers suado pela função `boxplot`. Esta função retorna uma lista com várias informações, inclusive quem são os outliers, e vc pode aproveitá-la para identificá-los.
- De toda forma, pense que este critério de tolerância deve levar em conta não necessariamente a magnitude absoluta do desvio, e sim a relativa ao tamanho total do crânio (erro de 1cm em um crânio rato é bem mais grave do que num de anta), ou à variação que há na medida (de onde se aplicaria o critério de outlier acima).

Trabalho final

Código da função

```
difD=function(x,dimD,lim,outlier=TRUE)
{
  collength=dim(x)[2]
  rowlength=dim(x)[1]
  distt=(matrix(NA,ncol=(collength-1)/dimD,nrow=rowlength/2)) # matriz vazia
  para ser preenchida com o cálculo das distâncias
  colnames(distt)=paste("Variavel",1:((collength-1)/dimD))
  a=0
  b=0
  if (dimD==2)
  {
    for (i in seq(1,rowlength, by=2))
    {
      a=a+1
      for (j in seq(2,collength, by=dimD))
      {
        b=b+1
        distt[a,b]=(((x[i,j]-x[i+1,j])^2)+((x[i,j+1]-x[i+1,j+1])^2))^(1/2) #
cálculo efetivo das distâncias
      }
      b=0
    }
  }
  if (dimD==3)
  {
    for (i in seq(1,rowlength, by=2))
    {
      a=a+1
      for (j in seq(2,collength, by=dimD))
      {
        b=b+1
        distt[a,b]=(((x[i,j]-x[i+1,j])^2)+((x[i,j+1]-
x[i+1,j+1])^2)+((x[i,j+2]-x[i+1,j+2])^2))^(1/2) # cálculo efetivo das
distâncias
      }
      b=0
    }
  }
  {
    resultado=list()
    especimes=(seq(1,(dim(x)[1]), by=2))
    r=1
    for (i in 1:(dim(distt)[1]))
    {
```

```

especime=c(x[especimes[i],1],x[(especimes[i])+1,1])
maiorquedif= t(as.data.frame(which(as.numeric(distt[i,])>lim)))
rownames(maiorquedif)="Variavel"
colnames(maiorquedif)=rep("",length((which(as.numeric(distt[i,])>lim))))
if (outlier==TRUE)
  {
  out=(boxplot.stats(distt[i,]))
  out=out$out
  out=t(as.data.frame(out))
  }
else
  {
  out=numeric(0)
  }
if (length(out)==0 & length(maiorquedif)==0)
  {
  temp=list(especime,cat("não"),cat("não"))
  }
else
  {
  if (length(out)==0)
    {
    temp=list(especime, maiorquedif,cat("não"))
    }
  if (length(maiorquedif)==0)
    {
    temp=list(especime, cat("não"),out)
    }
  if (length(out)!=0 & length(maiorquedif)!=0)
    {
    temp=list(especime,maiorquedif, out)
    }
  }
  resultado[[r]]=temp
  r=r+1
}
return(resultado)
}
}

```

Help da função

difD

package:unknown

R

Documentation

Description

Explicita os pares de amostras de morfometria geométrica no espaço da forma sem os efeitos do tamanho, da translação e da rotação ("shape space") para os quais há distâncias euclidianas entre cada ponto homólogo maiores que um

limite pré-determinado pelo usuário (apresenta inclusive quais são estes pontos).

Usage

```
difD = function(x, dimD, lim, outlier=TRUE)
```

Arguments

`x` data.frame. Na primeira coluna deve constar o nome das amostras. Nas demais colunas seguem as coordenadas dos pontos.

`dimD` informa se os dados coletados foram em duas ou três dimensões (aceita dados 2D e 3D).

`lim` valor limite utilizado pela função para determinar quais pares de amostra apresentam distâncias entre os pontos homólogos superiores ao valor determinado pelo usuário.

`outlier` variável lógica. Se TRUE, apresenta as distâncias atípicas de acordo com o critério para valores atípicos adotado nas análises de Box Plot.

Details:

Função utilizada para comparar, par a par, amostras de morfometria geométrica obtidas tanto em duas (x,y) quanto em três dimensões (x,y,z). Opcionalmente pode apresentar as distâncias atípicas de acordo com o critério para valores atípicos adotado nas análises de Box Plot.

Value:

A função retorna uma lista composta de listas minoritárias:

Para cada lista minoritária são apresentadas: 1) as amostras comparadas; 2) os pontos (variáveis) com valores para as distâncias superiores ao limite imposto; e 3) quando solicitado, os pontos com distâncias atípicas.

Warning:

No presente momento esta função só é eficaz para observações em que os efeitos do tamanho, da rotação e da translação já tenham sido removidos previamente à análise.

Note:

A função compara as amostras do primeiro par do data.frame, em seguida do segundo par e assim sucessivamente. Ou seja, esta função confronta os pontos das amostras das linhas 1 e 2 do data.frame, depois das linhas 3 e 4, 5 e 6, etc.

Em breve uma nova versão desta função (difDplus) lidará também com as questões do tamanho, da rotação e da translação nas observações.

Author(s):

Alex Hubbe, com sugestões de Paulo Inácio.

References:

See Also:**Examples:**

```
#criando dados fictícios de uma amostra de 4 observações com 11 pontos em
#três dimensões
nomes=paste("amostra",1:4,sep="")
am1=c(0,0,0.009,-0.007)
am2=c(0,0,0.068,0.042)
am3=c(0,0,-0.015,0.018)
am4=c(0,0,0.003,0.006)
am5=c(0,0,0.076,0.073)
am6=c(0,0,-0.008,0.003)
am7=c(0,0,-0.037,-0.038)
am8=c(0,0,0.038,0.037)
am9=c(0,0,0.082,0.083)
am10=c(0,0,0.009,0.009)
am11=c(0,0,5.068,0.068)
am12=c(0,0,-5.015,-0.015)

exemplo=data.frame(nomes,am1,am2,am3,am4,am5,am6,am7,am8,am9,am7,am8,am9,am7
,am8,am9,am7,am8,am9,am7,am8,am9,am7,am8,am9,am7,am8,am9,am10,am
11,am12, stringsAsFactors = FALSE)

### execução da função com a apresentação dos valores atípicos (primeiro
caso)e sem eles (segundo caso)
difD(exemplo,3,0.01)
difD(exemplo,3,0.0099, outlier=FALSE)
```

From:

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

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2010:alunos:trabalho_final:alexhubbe:start Last update: **2020/08/12 06:04**