

Joyce Prado



Sou bióloga, estudante de doutorado do programa de pós-graduação em Ecologia Aplicada (ESALQ/CENA), e atuo na área de sistemática e taxonomia de pequenos mamíferos.

exec

TRABALHO FINAL

Proposta A

Classicamente dados morfológicos são comparados visualmente na forma de diagramas Dice-Leraas, desenvolvido por Simpson (1961), esse procedimento em estudos de variação geográfica, permite avaliar diretamente, na geografia, a existência de descontinuidades significativas ou não, em amostras com distribuição normal. Esses diagramas são construídos com base nos intervalos de confiança de 95% da média, e a média + ou - o desvio padrão. A vantagem deste tipo de diagrama é que ele não só mostra a diferença entre as amostras, mas a regressão da variável medida. Assim minha proposta será elaborar uma função, na qual o objeto de entrada será um data frame, onde na primeira coluna estará as populações e nas colunas seguintes estarão os dados de medidas morfológicas para diferentes variáveis. O objeto de saída será uma composição de gráficos, onde cada variável terá um gráfico mostrando sua média por população, bem como seu intervalo de confiança de 95%, porém antes do cálculo das métricas a função realizará um teste para ver se a distribuição das variáveis é normal, uma vez que essa é uma das premissas dos diagramas Dice-Leraas.

Proposta B

Relações proporcionais entre as diferentes combinações de espécies são frequentemente utilizados nos estudos de taxonomia, e podem ser ilustrados através de diagramas de razão. Minha proposta será uma função que facilitará a construção desses diagramas, onde o objeto de entrada serão medidas morfométricas para diferentes grupos os quais se queiram comparar. Para cada variável serão calculados o valor absoluto da média, e + ou - duas vezes o desvio padrão da média, os quais serão convertidos em logaritmos. Em seguida, o logaritmo da média da amostra padrão (amostra a qual as outras serão comparadas) será subtraído dos logaritmos da média, e + ou - duas vezes o desvio padrão, da amostra comparativa. As medidas maiores do que da amostra padrão serão, representados por valores positivos, as medidas menores por valores negativos. Linhas contínuas ou

tracejadas em cada diagrama irão conectar as médias amostrais e barras horizontais representarão + ou - duas vezes o desvio padrão da média. A amostra que tiver as mesmas proporções que a amostra padrão será representada por valores médios em uma linha paralela à do padrão, independentemente do tamanho absoluto. Além disso, se os valores da amostra que está sendo comparada com o padrão é semelhante em tamanho absoluto, elas estarão juntas no diagrama. Se as proporções entre qualquer das dimensões medidas são semelhantes, as posições dos seus pontos em relação ao outro na escala horizontal será semelhante. O objeto de saída será um diagrama de razão com a comparação entre os diferentes grupos.

Joyce, a proposta A parece ser muito simples em termos de execução. Acho que você deveria fazer com que ela gerasse mais dados do que apenas os gráficos com as médias das respectivas variáveis morfométricas para cada população. Tenho algumas dúvidas: o que são as regressões geradas? Essas regressões farão parte do resultado final da função?

A segunda proposta é mais interessante em termos de resultado e execução, eu investiria nela. No entanto, tenho algumas dúvidas sobre os diagramas gerados pela função. Você tem exemplos? Quais seriam os argumentos da função gerada? Entre em contato para maiores esclarecimentos. —
[Glaucia Del-Rio](#)

Olá Glaucia, Muito Obrigada pelos comentários, escrevi um email para ti com o anexo do modelo do gráfico que pretendo fazer.

RESOLUÇÃO PROPOSTA B

Página de Ajuda

ratiodiagram

package: nenhum

R Documentation

~~Diagrama de Proporção~~

Description:

Ilustra as relações proporcionais entre as médias e o erro padrão de diferentes grupos.

Usage:

```
ratiodiagram (x, name="figura.tif", ngrupos, rmNA=TRUE, erro=TRUE)
```

Arguments:

`x` matriz ou `data frame` contendo na primeira coluna os grupos que serão comparados, e nas colunas seguintes `as` variáveis que serão utilizadas na comparação dos grupos.

`name` nome no qual a figura será exportada para a área de trabalho.

`ngrupos` número de grupos existentes no objeto `x`. 0 número de grupos existentes deve ser maior que um, senão não será executada a computação.

`na.rm` valor lógico que indica se os valores de NA devem ser removidos antes que computação prossiga.

`erro` indica se `as` barras do erro padrão da média devem ser plotadas no gráfico.

Details:

Relações proporcionais entre `as` médias das variáveis de diferentes combinações de grupos são frequentemente utilizados, e podem ser ilustradas através de diagramas de razão.

Essa função facilita a construção desses diagramas, sendo o objeto de entrada uma matriz de variáveis contínuas para diferentes grupos os quais se deseja comparar.

Para cada amostra são calculados o valor absoluto da média de cada variável e o valor destas médias $\pm$ duas vezes o erro padrão da média, e então estes três valores calculados são logaritmizados.

Em seguida, o logaritmo da média da amostra padrão, ou seja, aquela à qual `as` outras serão comparadas, é subtraído dos logaritmos da média e dos valores de média $\pm$ duas vezes o erro padrão de cada uma das outras amostras. As medidas maiores do que da amostra padrão são representados por valores positivos, `as` medidas menores por valores negativos.

Linhas irão conectar `as` médias das variáveis de cada amostra e barras horizontais representarão $\pm$ duas vezes o desvio padrão da média.

A logaritmização dos dados elimina a distorção causada pelas diferenças de escala entre `as` variáveis, permitindo que todas sejam comparadas simultaneamente e de maneira proporcional entre os grupos.

Quanto mais próximas `as` linhas de cada amostra, maior a similaridade entre `as` médias de suas variáveis. O objeto de saída é um diagrama de razão com a comparação entre os diferentes grupos.

Value:

A função `ratiodiagram` retorna o seguinte resultado:

Gráfico com `as` proporções das médias em imagem no formato TIFF

Warning:

Os grupos são organizados em ordem alfabética durante a computação e sempre o último grupo da lista será usado como a amostra padrão, à qual os outros grupos serão comparados.

Author(s):

Joyce Rodrigues do Prado

References:

MUSSER, G.G.; CARLETON, M.D.; BROTHERS, E.; GARDNER, A.L. *Systematic studies of Oryzomyine rodents (Muridae, Sigmodontinae): diagnoses and distributions of species formerly assigned to Oryzomys "capito"*. *Bulletin of the American Museum of Natural History*, New York, v. 236, p. 1-376, 1998.

SIMPSON, G.G. *Large Pleistocene Felines of North America*. *American Museum Novitates*, New York, v. 1136, p.1-27, 1941.

Examples:

```
#Para rodar os exemplos a seguir é necessário carregar o arquivo teste
dados<- read.table("teste.csv", header = TRUE, sep = ",")
ratiodiagram (teste)
ratiodiagram (teste, ngrupos, na.rm=FALSE)
ratiodiagram (teste, name="plot.tif", erro=FALSE)
```

Código da Função

```
# Disciplina: Uso da Linguagem R para Análise de Dados em Ecologia - 2014
# Função: Diagrama de Proporção
# Autor: Joyce Rodrigues do Prado

ratiodiagram<- function(x,name="figura.tif", ngrupos, rmNA=TRUE, erro=TRUE)
{
  dados<-x #Nomeia o objeto de entrada
  #Se o número de grupos for menor ou igual a 1 não é possível executar a
  função e aparece uma mensagem de erro.
  ngrupos<-nlevels(dados[,1])
  if (ngrupos<=1)
  {
    stop("Atenção! ngrupos deve ser maior que 1.")
  }
  #Se o número de grupos for maior que 1 a função executa normalmente.
  else
  {
    grupos<-levels(dados[,1]) # mostra quais são os grupos que serão
    comparados
    if (rmNA==FALSE)
    {
      stop("Atenção! Não é possível executar a função com NAs.")
    }
  }
}
```

```

else
  if (erro==TRUE)
  {
    media<- aggregate(dados[, -1], list(grupos=dados[, 1]), mean,
na.rm=T) #calcula a média de cada variável em cada grupo
    erro<- aggregate(dados[, -1], list(grupos=dados[, 1]),
function(x, na.rm = T) 2* sd(x, na.rm = T)/sqrt(length(x))) # calcula o erro
padrão da amostra por variável
    media_pluserro<-media[, -1]+erro[, -1] # soma o erro padrão à
média
    media_minorerro<-media[, -1]-erro[, -1]# diminui o erro padrão
da média
    logmedia<- log10(media[, -1]) #logaritmiza a média
    logmedia_pluserro<- log10(media_pluserro) #logaritmiza a
soma da média mais erro padrão
    logmedia_minorerro<- log10(media_minorerro)#logaritmiza a
diferença da média mais erro padrão
    comb1 <- expand.grid(x1=nrow(logmedia), x2=1:nrow(logmedia))
#Faz uma combinação dos grupos para calcular a diferença entre as médias dos
grupos
    XSub1 <- logmedia[comb1$x1,] - logmedia[comb1$x2,] #Calcula
a diferença entre as médias dos grupos
    rownames(XSub1) <- paste(comb1$x1, comb1$x2, sep="-") #Muda
os nomes das linhas
    transposta1<-t(XSub1) #transpõe a matriz pra fazer o gráfico
    comb2 <- expand.grid(x1=nrow(logmedia_pluserro),
x2=1:nrow(logmedia_pluserro)) #Faz uma combinação dos grupos para calcular
a diferença entre a soma da média e do erro padrão dos grupos
    XSub2 <- logmedia_pluserro[comb2$x1,] -
logmedia_pluserro[comb2$x2,] #Calcula a diferença entre a soma da média e do
erro padrão dos grupos
    rownames(XSub2) <- paste(comb2$x1, comb2$x2, sep="-") #Muda
os nomes das linhas
    transposta2<-t(XSub2)#transpõe a matriz pra fazer o gráfico
    comb3 <- expand.grid(x1=nrow(logmedia_minorerro),
x2=1:nrow(logmedia_minorerro))#Faz uma combinação dos grupos para calcular a
diferença entre a diferença da média e do erro padrão dos grupos
    XSub3 <- logmedia_minorerro[comb3$x1,] -
logmedia_minorerro[comb3$x2,] #Calcula a diferença entre a diferença da
média e do erro padrão dos grupos
    rownames(XSub3) <- paste(comb3$x1, comb3$x2, sep="-") #Muda
os nomes das linhas
    transposta3<-t(XSub3) #transpõe a matriz pra fazer o gráfico
    nomeslinhas<-rownames(transposta1) #cria um objeto com o
nome das variáveis
    i<- length(nomeslinhas) # calcula número de variáveis
    max2<- max(transposta2, na.rm =TRUE)#achando o maior valor
da matriz transposta2 para usar no eixo y do gráfico
    max1<- max(transposta1, na.rm =TRUE)#achando o maior valor
da matriz transposta1 para usar no eixo y do gráfico
    max3<- max(transposta3, na.rm =TRUE)#achando o maior valor

```

```
da matriz transposta3 para usar no eixo y do gráfico
    min1 <- min(transposta1, na.rm =TRUE)#achando o menor valor
da matriz transposta1 para usar no eixo y do gráfico
    min2<- min(transposta2, na.rm =TRUE)#achando o menor valor
da matriz transposta2 para usar no eixo y do gráfico
    min3 <- min(transposta3, na.rm =TRUE)#achando o menor valor
da matriz transposta3 para usar no eixo y do gráfico
    max<-max(c(max2,max1,max3))#achando o maior valor de todas
as transpostas para usar no eixo y do gráfico
    min<- min(c(min1,min2,min3))#achando o menor valor de todas
as transpostas para usar no eixo y do gráfico
#Abre um dispositivo tiff para salvar o gráfico.
    tiff(file= name, width=40, height=15, unit="cm", res=300)
#Faz o gráfico
    par (family="serif")#fonte
    par(cex= 1)#tamanho dos simbolos
    par (cex.axis=1)#tamanho caracteres dos eixos
    par(cex.lab=1.2)# tamanho legenda dos eixos
    par (tcl=0.3)#posição dos tickmarks
    par(mgp=c(1.3,0.1,0))#distância da legenda em relação ao
eixo
    par(bty="l")#tipo de caixa gráfica
    grafico<-matplot(transposta1, lty= c(1,2,3,4,5,6),tcl=
-0.2,type="o",xaxt="n", pch=c(0,1,2,3,4,5,6,7,8,9), col="black", xlab="",
ylab="", ylim= c(min,max))
    axis (side=1, at= c(1:i),label= nomeslinhas,tck=0)#editando
eixo x
    legend("topright", legend = grupos, bty="n", lty=
c(1,2,3,4,5,6), pch=c(0,1,2,3,4,5,6,7,8,9))#colocando legenda
    segments (x0=c(1:i), x1=c(1:i), y0= transposta2, y1=
transposta3)#colocando barra de erros
    segments (x0= c(1:i-0.1), x1= c(1:i+0.1), transposta2,
transposta2)#colocando barra de erros
    segments (x0= c(1:i-0.1), x1= c(1:i+0.1), transposta3,
transposta3)#colocando barra de erros
#Fecha o dispositivo tiff.
    dev.off()#fechando dispositivo
# Retorna o gráfico de proporções
    return("o gráfico foi salvo na pasta de trabalho")
}
else
{
    media<- aggregate(dados[, -1], list(grupos=dados[, 1]), mean,
na.rm=T) #calcula a média de cada variável em cada grupo
    logmedia<- log10(media[, -1]) #logaritmiza a média
    comb1 <- expand.grid(x1=nrow(logmedia), x2=1:nrow(logmedia))
#Faz uma combinação dos grupos para calcular a diferença entre as médias dos
grupos
    XSub1 <- logmedia[comb1$x1,] - logmedia[comb1$x2,] #Calcula
a diferença entre as médias dos grupos
```

```
rownames(XSub1) <- paste(combl$x1, combl$x2, sep="-") #Muda
os nomes das linhas
transpostal<-t(XSub1) #transpõe a matriz pra fazer o gráfico
nomeslinhas<-rownames(transpostal) #cria um objeto com o
nome das variáveis
i<- length(nomeslinhas) # calcula número de variáveis

#Abre um dispositivo tiff para salvar o gráfico.
tiff(file= name, width=40, height=15, unit="cm", res=300)
#Faz o gráfico
par (family="serif")
par(cex= 1)
par (cex.axis=1)
par(cex.lab=1.2)
par (tcl=0.3)
par(mgp=c(1.3,0.1,0))
par(bty="l")
grafico<-matplot(transpostal, lty= c(1,2,3,4,5,6),tcl=
-0.2,type="o",xaxt="n", pch=c(0,1,2,3,4,5,6,7,8,9), col="black", xlab="",
ylab="")
axis (side=1, at= c(1:i),label= nomeslinhas,tck=0)
legend("topright", legend = grupos, bty="n", lty=
c(1,2,3,4,5,6), pch=c(0,1,2,3,4,5,6,7,8,9))
#Fecha o dispositivo tiff.
dev.off()
# Retorna o gráfico de proporções
return("o gráfico foi salvo na pasta de trabalho")
}
}
}
```

[codigo_da_funcao.r](#) [teste.csv](#)

From:

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

Permanent link:

http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2014:alunos:trabalho_final:joyce.prado:start

Last update: 2020/09/23 20:16

