

# Ramiro Campelo Araújo Ribeiro

Mestrando em Ecologia, Instituto de Biociências, USP. Laboratório de Estudos Evolutivos Humanos (LEEh), divisão de Ecologia Humana.

exec

## Propostas de trabalho final

### Contextualização:

A partir de abordagens de orientação interdisciplinar, diversas linhas de pesquisa situadas na intersecção entre as ciências naturais e humanas almejam não só uma permuta metodológica e teórica entre disciplinas, mas também práticas norteadas por perguntas e orientações teóricas em comum. Tal integração é certamente um dos maiores desafios do empenho interdisciplinar. As técnicas quantitativas empregadas nessas áreas buscam um refinamento das metodologias e fortalecem tal integração. Produzir valores confiáveis e comparáveis a partir de dados qualitativos menos tangíveis é o primeiro passo desse processo.

Em ambos polos acadêmicos, uma ferramenta comum para uma aproximação preliminar entre dados qualitativos e quantitativos é representada pela aplicação de diversos tipos de índices. Na etnobiologia e etnoecologia, os “Índices de saliência cultural relativa” (exemplos na tabela 5, pág.208 <http://lib-ojs3.lib.sfu.ca:8114/index.php/era/article/download/130/115>) são medidas quantitativas utilizadas para transformar conceitos complexos e multidimensionais como o “valor de uso” em valores ou escalas numéricas padronizadas e comparáveis, provendo dados passíveis de teste de hipóteses, validação estatística e análises comparativas. Portanto, ao avaliar a correspondência de dados transformados por índices que avaliam diferentes facetas de uma mesma dimensão do conhecimento etnoecológico, podemos fortalecer o potencial de comparações trans-culturais, e consequentemente um aprofundamento da compreensão das relações entre populações humanas, a paisagem vegetal e as instâncias práticas, utilitárias e simbólicas do uso de plantas.

### A função:

#### Plano A

O papel da função será automatizar uma rotina de análises exploratórias, baseada na reorganização, padronização e extração de dados de um dataframe para subseqüentes cálculos de alguns dos diversos índices supracitados, que serão, então, comparados por meio da correlação de Pearson e análise de componentes principais (PCA), e suas correspondências estabelecidas. Tal processo almeja proporcionar um resultado que facilite a subseqüente realização de testes de hipóteses, assim como outros procedimentos analíticos menos triviais.

### Entrada:

Dataframe contendo variáveis qualitativas – binárias e nominais – e quantitativas – escalas ordinais de frequências, abundâncias e rankings.

## Argumentos:

- Entrada do objeto de dados;
- Argumento lógico de remoção de NA's;
- Definição dos tipos de índices a serem calculados;
- Definição da forma de comparação dos resultados dos índices (Pearson ou PCA)

## Passos:

1. Reorganização das variáveis em vetores correspondentes ao tipo de variável;
2. Realização dos cálculos dos índices;
3. Realização da correlação de Pearson;
4. Realização do PCA;
5. Produção de tabelas com os resultados;

## Objetos de saída:

- Tabelas contendo os resultados dos índices individualmente para cada evento (observação);
- Tabelas contendo o resultado das comparações pela correlação de Pearson e PCA;

## Exemplificando:

Dada a especificidade dos índices em questão, a função será utilizada em investigações sobre o conhecimento de populações tradicionais acerca do repertório etnobotânico e etnoecológico local. Dessa forma, as variáveis consistem em informações obtidas por diversos métodos provenientes do componente etnográfico do trabalho, e incluem as categorias locais da paisagem, as plantas que constituem o repertório (as identidades locais e científicas), as categorias de utilização, detalhes da forma de preparo e aplicações, estruturas (em termos anatômicos), ocorrência desses elementos do repertório nas categorias locais da paisagem, e a frequência (em uma escala ordinal que parte da percepção dos informantes) dessas ocorrências.

## Plano B

Como alternativa, proponho a realização de um resumo visual básico, com gráficos elementares provenientes de uma análise exploratória simples. Entretanto, Uma vez que existem inúmeros resumos numéricos e gráficos possíveis, o objetivo do Plano B consistiria em analisar as mesmas variáveis citadas no Plano A, tornando possível o retorno dessa ampla gama de representações gráficas e numéricas a partir da utilização da filogenia como um *proxy*, ou seja, o plano implicaria em uma forma de plotar o APG III (ação provavelmente executada por alguma função existente) de forma que os resumos visuais possam ser requisitados a partir da escolha de diferentes níveis hierárquicos. A saída seria composta pelo resumo visual das variáveis escolhidas mediante a escolha de um nível de resolução taxonômica. Assim, os padrões de uso poderiam ser observados em contraposição a dados filogenéticos das plantas pertencentes ao repertório etnobotânico estudado.

Ola, Ramiro! A proposta A parece interessante e bem dimensionada para a função final. Faltaram detalhes pra avaliar melhor o plano B, mas acho que não será necessário se você optar pelo plano A.

—Leo

Concordo com o Leo,  **siga o plano A**. Na ecologia quando temos variáveis de naturezas diferentes (contínuas, categóricas, binárias) em geral aplicamos o índice de Gower. Veja se ele é útil para o seu caso - veja Lengender & Legender (Numerical Ecology, pag.254). Não vejo como aplicar um PCA nesse caso, pode ser ignorância minha, mas achei que fosse necessário uma escala contínua e padronizada de cada variável.

## Código da função

```
ethnobot.indxs<-function(X){
dados<-X

if(class(dados)!="data.frame"){

  stop("Objeto de entrada de dados deve ser um dataframe")
}else if (is.null(dados$informante)){
  stop("O dataframe deve conter uma coluna chamada 'informante' ")
}else if (is.null(dados$etnoespecie)){
  stop("O dataframe deve conter uma coluna chamada 'etnoespecie' ")
}else if(is.null(dados$especie)){
  stop("O dataframe deve conter uma coluna chamada 'especie' ")
}else if(is.null(dados$familia)){
  stop("O dataframe deve conter uma coluna chamada 'familia' ")
}else if(is.null(dados$parte)){
  stop(" O dataframe deve conter uma coluna chamada 'partes' ")
}else if(is.null(dados$ordem)){
  stop("O dataframe deve conter uma coluna chamada 'ordem' ")
}else if(is.null(dados$comunidade)){
  stop("O dataframe deve conter uma coluna chamada 'comunidade' ")
}
df1<-dados[which(dados[,1]=="a"),-1]
df2<-dados[which(dados[,1]=="b"),-1]

categoricas <- names(df1)[1:6]

numericas <- names(df1)[7:length(df1)]

d1 = melt(df1, id.vars = categoricas)

d1 = rename(d1, replace=c("variable"= "usos"))
```

```
d1=rename(d1,replace=c("value"="valores"))

RU1 <- ddply(d1, .(etnoespecie), numcolwise(sum))

RU.pp1 <- ddply(d1, .(etnoespecie, parte), numcolwise(sum))

df.SU1<-df1[, -c(1,2,4:6)]
df.SU1<-aggregate(x=df.SU1[, -1],by=list(name=df.SU1[,1]),FUN=sum)
df.SU1[, -1][df.SU1[, -1]!=0]<-1
SU1<-data.frame(etnoespecie=df.SU1[,1],SU1=apply(df.SU1[, -1],1,sum))

RU.pp1$PPV=0
for(i in 1:nrow(RU.pp1)){
  RU.pp1$PPV[i]<-RU.pp1[,3][i]/RU1[,2][RU.pp1[,1][i]==RU1[,1]]
}

df.SU.pp1<-df1[, -c(1,4:6)]
df.SU.pp1<-aggregate(x=df.SU.pp1[, -
c(1,2)],by=list(parte=df.SU.pp1[,1],etnoespecie=df.SU.pp1[,2]),FUN=sum)
df.SU.pp1[, -c(1,2)][df.SU.pp1[, -c(1,2)]!=0]<-1
SU.pp1<-
data.frame(parte=df.SU.pp1[,1],etnoespecie=df.SU.pp1[,2],SU.pp=apply(df.SU.p
pp1[, -c(1,2)],1,FUN=sum))

OUV1<-
data.frame(etnoespecie=df.SU.pp1[,2],parte=df.SU.pp1[,1],OUV=(SU.pp1[,3]/RU.
pp1[,3])*RU.pp1$PPV)

categoricas <- names(df2)[1:6]

numericas <- names(df2)[7:length(df2)]

d.2 = melt(df2, id.vars = categoricas)

d.2 = rename(d.2, replace=c("variable"= "usos"))

d.2=rename(d.2,replace=c("value"="valores"))

RU2 <- ddply(d.2, .(etnoespecie), numcolwise(sum))

RU.pp2 <- ddply(d.2, .(etnoespecie, parte), numcolwise(sum))

df.SU2<-df2[, -c(1,2,4:6)]
df.SU2<-aggregate(x=df.SU2[, -1],by=list(name=df.SU2[,1]),FUN=sum)
df.SU2[, -1][df.SU2[, -1]!=0]<-1
SU2<-data.frame(etnoespecie=df.SU2[,1],SU=apply(df.SU2[, -1],1,sum))

RU.pp2$PPV=0
for(i in 1:nrow(RU.pp2)){
  RU.pp2$PPV[i]<-RU.pp2[,3][i]/RU2[,2][RU.pp2[,1][i]==RU2[,1]]
}
```

```
}  
  
df.SU.pp2<-df2[, -c(1,4:6)]  
df.SU.pp2<-aggregate(x=df.SU.pp2[, -  
c(1,2)],by=list(parte=df.SU.pp2[,1],etnoespecie=df.SU.pp2[,2]),FUN=sum)  
df.SU.pp2[, -c(1,2)][df.SU.pp2[, -c(1,2)]!=0]<-1  
SU.pp2<-  
data.frame(parte=df.SU.pp2[,1],etnoespecie=df.SU.pp2[,2],SU.pp2=apply(df.SU.  
pp2[, -c(1,2)],1,FUN=sum))  
  
OUV2<-  
data.frame(etnoespecie=df.SU.pp2[,2],parte=df.SU.pp2[,1],OUV=(SU.pp2[,3]/RU.  
pp2[,3])*RU.pp2$PPV)  
  
##Correlação de Pearson para as duas comunidades  
x11()  
plot(OUV1[,3]~OUV2[,3],main="Regressão Linear",xlab="OUV comunidade  
2",ylab="OUV comunidade 1")  
abline(lm(OUV1[,3]~OUV2[,3]))  
pearson<-cor(OUV1[,3],OUV2[,3])  
resultado<- list("pearson"=pearson,"OUV comunidade a"= OUV1,"OUV comunidade  
b"= OUV2)  
return(resultado)  
}
```

From:

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

Permanent link:

[http://ecor.ib.usp.br/doku.php?id=05\\_curso\\_antigo:r2015:alunos:trabalho\\_final:ramiroaraujo:start](http://ecor.ib.usp.br/doku.php?id=05_curso_antigo:r2015:alunos:trabalho_final:ramiroaraujo:start)



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