

Partição de variação: uma espécie


Os três parâmetros estruturais básicos das comunidades são a riqueza, composição e abundância (de espécies, tipos funcionais, linhagens ...). Uma das abordagens da ecologia de comunidades é buscar regularidades nesses parâmetros, e propor explicações para eles. Neste roteiro vamos aprender uma das ferramentas estatísticas usadas para isto, chamada *partição de variação*.

Podemos tentar entender as comunidades analisando cada uma das populações de cada espécie presente ou podemos tentar entender o conjunto de espécies como um todo. Aqui vamos analisar a estrutura espacial da abundância de uma espécie fictícia. Começamos com apenas uma espécie porque isso facilita entender os cálculos e suas interpretações. No roteiro seguinte mostraremos como a partição pode ser generalizada para descrever padrões conjuntos de abundância de muitas espécies.

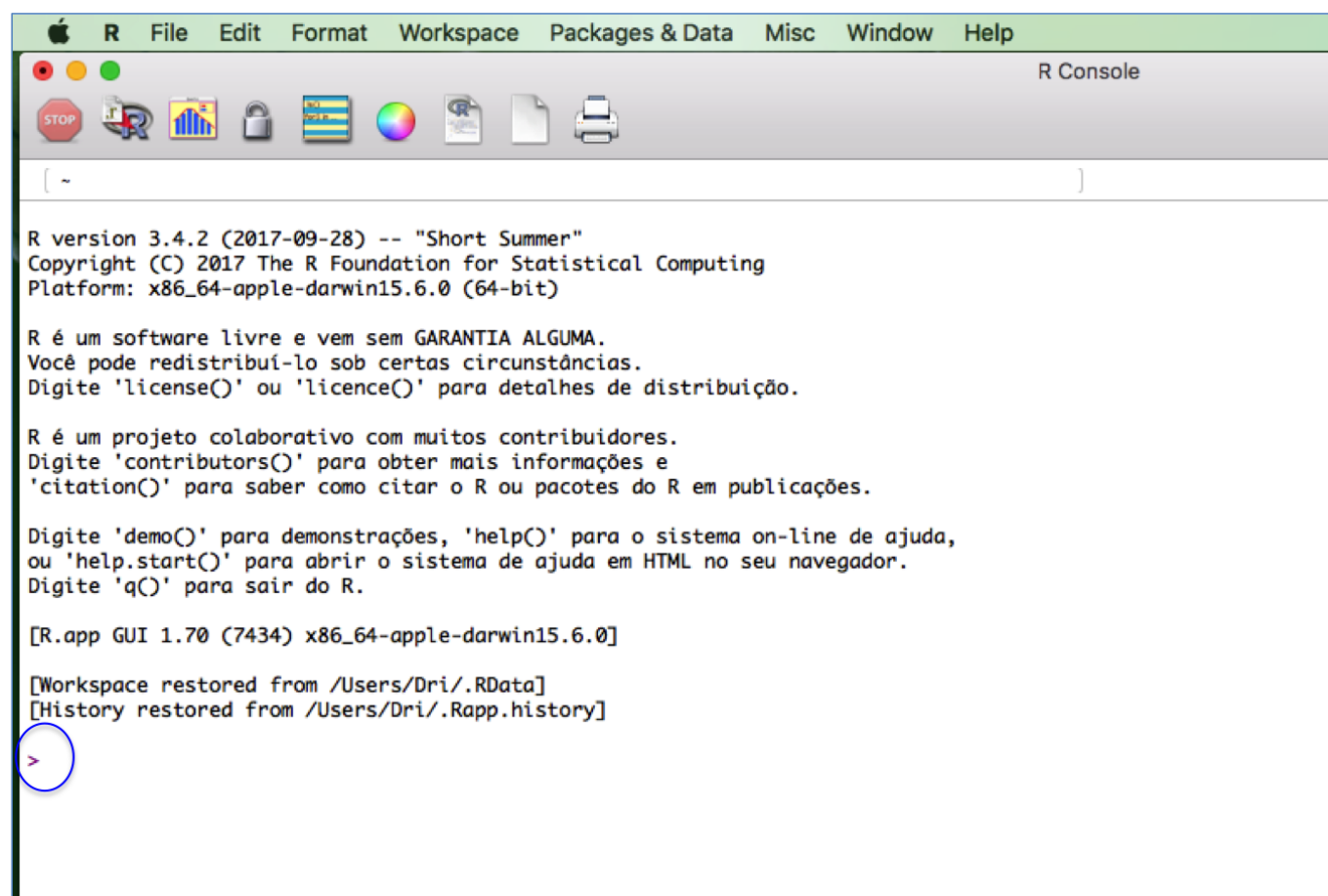
Preparação para o exercício

Para começar, escolha uma pasta para você guardar os arquivos e resultados deste roteiro em seu computador. Copie para essa pasta o arquivo com os dados que vamos usar:

- roteiro1b.csv

Em seguida, abra o programa R, clicando no ícone  que está na área de trabalho do seu computador.

Se tudo deu certo até aqui, abrirá uma janela do R como essa:



Já com a janela do programa R aberto, o próximo passo será mudar o diretório de trabalho para aquela pasta que você acabou de criar. Com isso será mais fácil importar os dados dos arquivos ".csv" para dentro do ambiente R.

A mudança de diretório deve ser feita da seguinte forma:

- Abra o Menu "Arquivo" (ou "File");
- Selecione "Mudar dir" (ou "Change dir");
- Escolha a sua pasta na janela que abrir.

[Obs. Para Mac, essa opção está no Menu "Misc" e a opção é "Change working dir"]

Para checar se você está na pasta correta, copie e cole o comando abaixo na linha de comando do R. Atenção: O comando deve ser colado na frente do símbolo ">", circundado em azul na imagem anterior. Este símbolo indica o início da linha de comando ou "prompt", onde você deve escrever comandos para o R.

```
getwd()
```

Após colar, aperte a tecla "enter" e veja se o R retorna o nome da sua pasta. Se sim, ótimo. Se não, chame um monitor ou professor.

Importando os dados para o R

Agora vamos importar para o R os dados que você gravou em seu diretório. Para isso copie o comando abaixo, cole na linha de comando do R e pressione “enter”:

```
dados <- read.csv("roteiro1b.csv")
```

Se não houve nenhuma mensagem de erro agora você tem no R uma tabela com 100 linhas e três colunas, que explicaremos a seguir. Se quiser verificar se a tabela foi importada, você pode usar os comandos seguintes para mostrar suas primeiras e últimas linhas:

```
head(dados)  
tail(dados)
```

Em busca de uma estrutura

A abundância de uma espécie pode variar muito, no tempo e no espaço. A pergunta fundamental aqui é o que causa esta variação. A partição de variância é um método para avaliar o quanto da variação de uma quantidade pode ser atribuída à variação de uma outra quantidade. Usamos este método para avaliar se é plausível que a abundância de uma planta em uma área varie em função da variação de um fator ambiental, por exemplo.

Um padrão

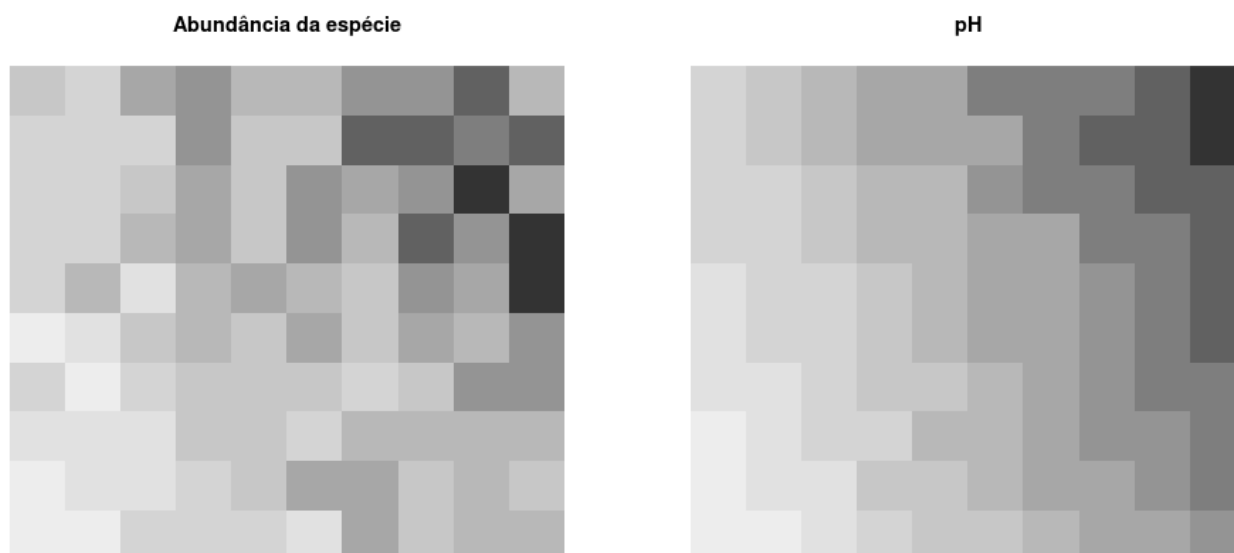
Vamos aos nossos dados, para deixar mais concreto. A figura abaixo representa uma parcela permanente, que foi dividida em 100 quadrados. Em cada quadrado foram contados todos os indivíduos de *Non illum*, uma erva daninha invasora. A abundância da planta por quadrado vai de dois indivíduos (cinza claro) até 41 indivíduos (cor preta).

Abundância da espécie



Uma hipótese: variação ambiental

Percebendo a concentração de plantas no canto superior direito da parcela, uma ecóloga propôs a hipótese de que a planta tem uma associação com solos menos ácidos. Ela chegou a esta hipótese porque havia medido o pH do solo superficial em cada parcela, o que usou para fazer o mapa a seguir (à direita). Neste mapa, cores claras são valores mais baixos de pH e cores escuras valores altos.



Parece razoável? Para avaliar melhor isto vamos fazer um gráfico da relação entre a abundância da planta e o valor do pH em cada parcela. Estes dados já estão na planilha que importamos para o R na seção anterior, então temos só que dar o comando que cria o gráfico:

```
plot(abund ~ pH, data = dados)
```

O gráfico sugere que a abundância da planta é proporcional ao pH. Se isso é verdade, deve haver uma linha reta que aproxime bem esta relação. A **regressão linear** é uma técnica estatística para encontrar esta reta. Execute no R o comando para ajustar a regressão linear da abundância em função do pH:

```
regr.pH <- lm(abund ~ pH, data = dados)
```

E adicione a linha da regressão ao seu gráfico com

```
abline(regr.pH)
```

A variação explicada pela hipótese

A regressão linear nos retorna muita informação importante sobre a relação que hipotetizamos. A que mais nos interessa agora é o **coeficiente de determinação**, expresso pela sigla R^2 . Este coeficiente expressa a proporção da variação total da variável resposta (a abundância, no caso) que é “capturada” ou “explicada” pela variação da variável explanatória (o pH, no caso).

Para obter o R^2 da regressão que ajustamos acima basta o comando¹⁾:

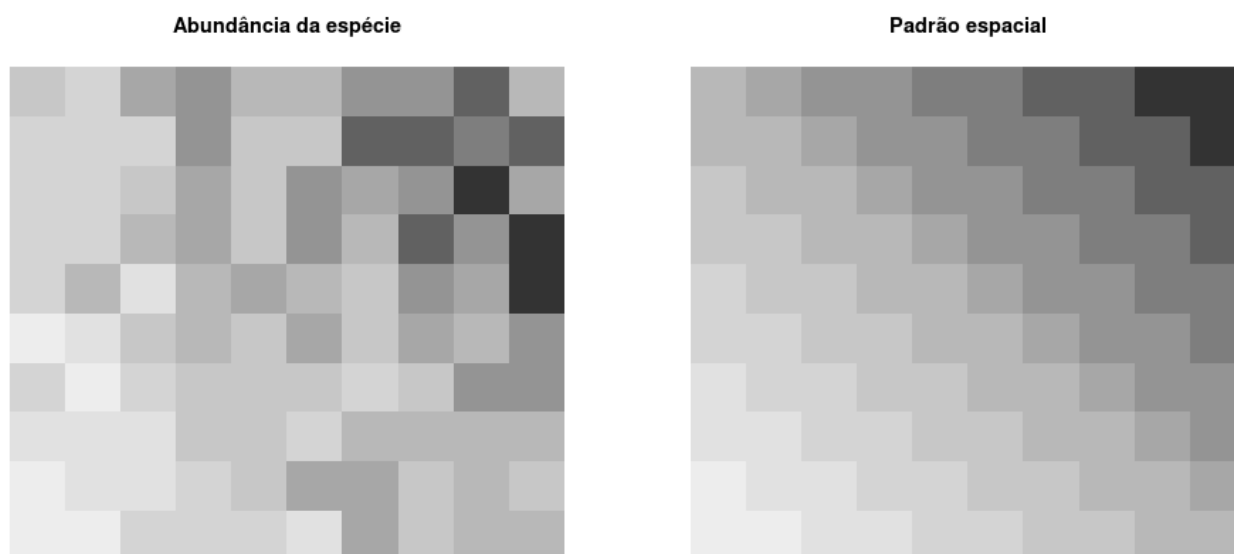
```
(X<-summary(regr.pH)$r.squared)
```

O que mostra que 68,6% da variação da abundância da espécie é explicada por uma relação linear com o pH do solo superficial.

Extra: Se ainda não fez, quando tiver tempo e quiser entender os cálculos do R^2 passo a passo experimente o [roteiro sobre coeficiente de determinação](#).

Outra hipótese: limitação à dispersão

Uma outra ecóloga sabia que a invasão por *Non illum* era recente e havia começado pelo canto superior direito da parcela por sementes trazidas acidentalmente por tucanos, que desapareceram em seguida ²⁾. A ecóloga então propôs que quadrados mais distantes do ponto de entrada tiveram menores chances de serem colonizados. Essa hipótese propõe que a variação da abundância desta planta observada na parcela é resultado apenas da **limitação à dispersão**. A ecóloga criou então um mapa para expressar sua hipótese. Neste mapa (à direita) os quadrados com maior chance de receber sementes estão em tons mais escuros:



Esse mapa é construído atribuindo um valor a cada quadrado, que deve expressar um índice de abundância relativa da espécie de acordo com o esperado pela hipótese proposta. No caso, este índice pode ser interpretado como proporcional à probabilidade de estabelecimento de uma semente. Uma hipótese de limitação à dispersão é expressa pelo arranjo espacial de áreas com alta e baixa probabilidade de estabelecimento, de acordo com algum modelo de dispersão ou estabelecimento. Neste caso o modelo é bem simples: quanto mais distante do vértice superior direito, menor a probabilidade de estabelecimento. Por isso vamos chamar o índice usado para construir o mapa de **variável de estrutura espacial**. Os valores dessa variável também já estão no arquivo que importamos para o R anteriormente.

Compare o mapa da direita com o mapa de variação da abundância da planta (esquerda). Será que a limitação à dispersão proposta é uma boa hipótese? Para ajudar, faça o gráfico da abundância em

função da variável de estrutura espacial com o comando abaixo:

```
plot(abund ~ space, data= dados, xlab = "Variável de estrutura espacial",  
     ylab= "Abundância de N. illum")
```

Teste da hipótese de limitação à dispersão

Para testar esta nova hipótese ajustamos a regressão linear da abundância em função da variável espacial:

```
regr.space <- lm(abund ~ space, data = dados)
```

E podemos adicionar a linha da regressão ao gráfico:

```
abline(regr.space)
```

Variação explicada pela hipótese de limitação à dispersão

Obtemos a fração da variação da abundância explicada pela variável espacial com o comando que já conhecemos, aplicado ao objeto que guarda os resultados da regressão:

```
(W <- summary(regr.space)$r.squared)
```

Partição da variação entre ambiente e espaço

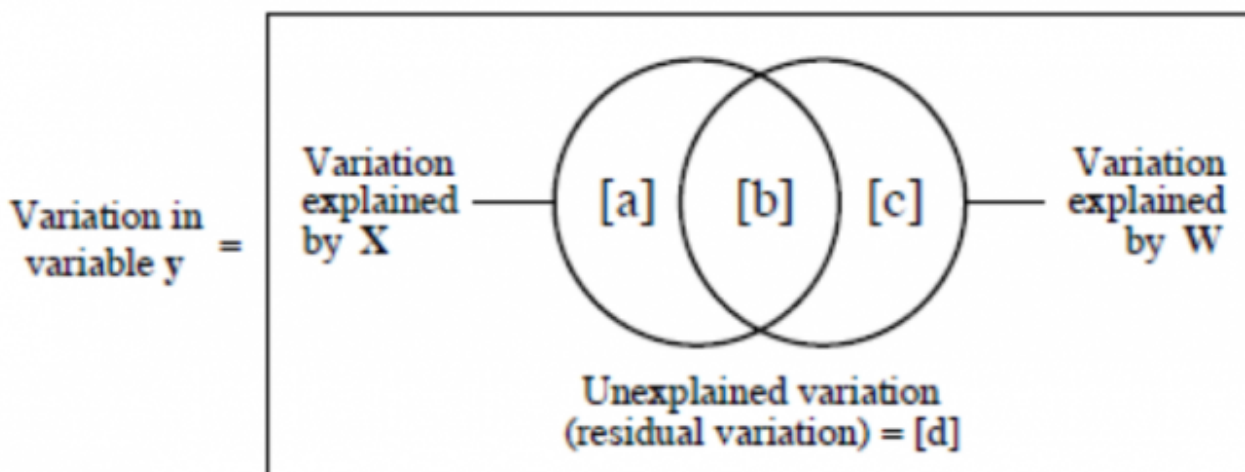
Resumindo o que encontramos até aqui: uma variável ambiental (pH) explica 68,6% da variação da abundância da espécie. Porém, uma variável de estrutura espacial explica 65,5% desta mesma variação. Como é possível que a soma dos percentuais de variação explicada passem de 100% ? 🤔

A resposta é que há uma relação entre o pH e a variável de estrutura espacial:

```
plot(pH ~ space, data = dados, xlab = "Variável de estrutura espacial")
```

O que significa que **a variável ambiental tem uma estrutura espacial**, que é bem descrita pela estrutura espacial proposta pela segunda hipótese. Assim, a variação explicada pelo pH contém uma parte de efeito espacial. Da mesma forma, uma parte da variação explicada pela variável espacial contém efeito do pH.

A esta parte da variação compartilhada chamamos **variação ambiental estruturada no espaço**. No esquema abaixo esta parte da variação explicada é o componente b .



O círculo da esquerda representa a variação explicada por uma variável X , no caso a variável ambiental. Vemos que a variação total explicada por X é a soma do componente compartilhado b com o componente a que chamamos **fração de variância explicada apenas pelo ambiente**.

Da mesma forma, o círculo da direita representa a variação explicada por uma segunda variável W , no caso o espaço. Vemos que W é formado pelo componente b (compartilhado com a outra variável) somado a uma outra fração c , que chamamos de **fração de variância explicada apenas pelo espaço**.

Por fim, a variação total é representada pelo retângulo externo. A parte não coberta pelos círculos é a variação que não é explicada por nenhuma variável, indicada por d .

Varição não explicada

A variação não explicada por nenhuma variável é obtida da regressão que inclui todas as variáveis:

```
regr.tudo <- lm(abund ~ pH + space, data = dados)
```

De onde obtemos a variação explicada pelo pH e a variável espacial juntas:

```
summary(regr.tudo)$r.squared
```

Esta quantidade corresponde a $a+b+c$ no diagrama acima, que vamos chamar de XW . Sabendo que

$$a+b+c+d = 1$$

temos então que

$$d = 1 - XW$$

Assim, obtemos d subtraindo o valor acima de um:


```
(d <- 1 - summary(regr.tudo)$r.squared)
```

Juntando as peças

Agora temos todas as peças que precisamos para calcular como a variação da abundância de *Non illum* pode ser particionada entre os componentes do diagrama acima:

$$\left\{ \begin{array}{l} a + b = X \\ c + b = W \\ a + b + c = 1 - d \end{array} \right.$$

O que implica em

$$d + X + W - 1 = b$$

Usando os valores que já calculamos e guardamos no R temos então:

Fração devido a espaço e pH (b)

```
(b <- X + W + d - 1)
```

Fração devido apenas a pH (a)

```
(a <- X - b)
```

Fração devido apenas a espaço (c)

```
(c <- W - b)
```

Fração não explicada (d)

```
d
```

Conclusão

A pergunta sobre a vida, o universo e tudo mais

Qual hipótese tem mais apoio dos resultados?

1)

com este comando também estamos guardando o valor obtido em uma objeto chamado X

2)

ao que parece tucanos e patos amarelos dispersam apenas uma vez as sementes, pois morrem envenenados por elas

Last
update:
2021/11/17 21:58 ecovirt:roteiro:particao:particao_univar http://ecovirtual.ib.usp.br/doku.php?id=ecovirt:roteiro:particao:particao_univar&rev=1637193511

From:

<http://ecovirtual.ib.usp.br/> -

Permanent link:

http://ecovirtual.ib.usp.br/doku.php?id=ecovirt:roteiro:particao:particao_univar&rev=1637193511



Last update: **2021/11/17 21:58**