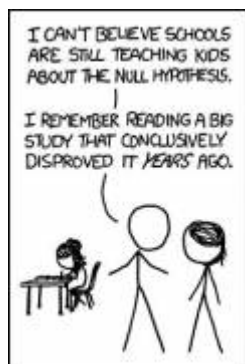


Teste de Hipóteses



O objetivo desse roteiro é auxiliar na compreensão de dois conceitos fundamentais do teste de hipótese estatístico. O primeiro é entender o significado do **p**, ou a probabilidade de rejeitar a hipótese nula quando ela não deveria ser rejeitada. Normalmente definimos um valor crítico baixo para essa probabilidade, dado que não queremos cometer esse erro. Convencionou-se, em muitas áreas da ciência, definir esse valor crítico em 0.05 ou 5%. O segundo objetivo desse roteiro é definir quando devemos aplicar um teste unicaudal ou bicaudal e como isso se relaciona à nossa hipótese estatística e à nossa hipótese biológica. É muito importante que, além de fazer o roteiro, você discuta com colegas e professores os conceitos apresentados.

Instalando o Rsampling

Para rodar esse roteiro utilizaremos duas ferramentas desenvolvidas para o ensino de estatística desenvolvida por pessoas ligadas ao departamento de ecologia do IBUSP. O primeiro é o pacote **Rsampling** que contém instrumentação de permutação e aleatorização para a linguagem R. O outro é uma interface gráfica (**Rsampling-shiny**) para usar as ferramentas deste pacote no seu navegador. Não há necessidade de conhecimento prévio em R para fazer esse roteiro.

Rsampling-shiny

O primeiro passo é baixar a interface gráfica e descompactar o arquivo em um diretório conhecido no seu computador. Sugerimos que crie uma pasta para a disciplina “planeco” dentro da pasta de documentos do computador que está trabalhando.

Baixe, para essa pasta, o arquivo “source code” [zip](#) ou [tar.gz](#)¹⁾ da versão mais recente (1.6.1- Small bugfix) do Rsampling-shiny no link:

- [Versões Rsampling-shiny](#)

Descompacte o arquivo na mesma pasta. Obs. depois de descompactar e testar o funcionamento (com os passos abaixo), você pode deletar o arquivo .zip ou .tar.gz

Usando o R

Abra o R no seu computador e mude o diretório de trabalho para o diretório (*i.e.* a pasta) que você criou, usando o menu **Arquivo > mudar dir....**

Instalando os pacotes

Em seguida copie e cole a seguinte linha de comando no terminal do R para instalar os pacotes associados:

```
install.packages(c("Rsampling", "shiny", "PerformanceAnalytics"))
```

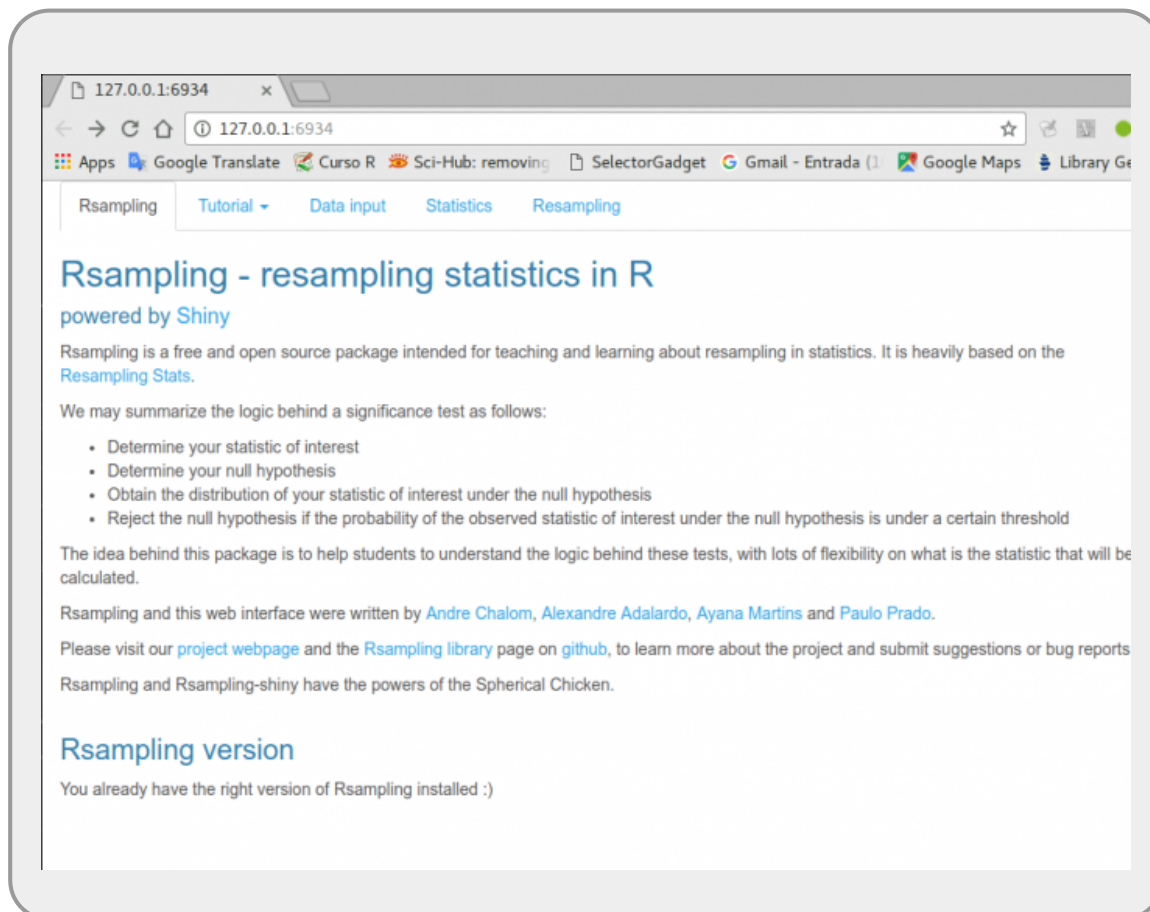
A seguinte mensagem irá aparecer no terminal do R “— Please select a CRAN mirror for use in this session —” e uma janela irá se abrir para que selecione o repositório mais próximo. Selecione **Brazil(SP 1)** Aguarde a instalação dos pacotes finalizar.

Abrindo a interface gráfica

Agora digite as seguintes linhas de comando no terminal do R. A primeira linha irá ativar o pacote shiny que acabou de instalar e a segunda irá abrir a interface Rsampling-shiny no seu navegador, a partir do nome do diretório que descompactou. Sempre que quiser abrir a interface, deve digitar essas duas linhas a partir do diretório da disciplina, não há necessidade de instalar os pacotes novamente.

```
library(shiny)  
runApp("Rsampling-shiny-1.6.1")
```

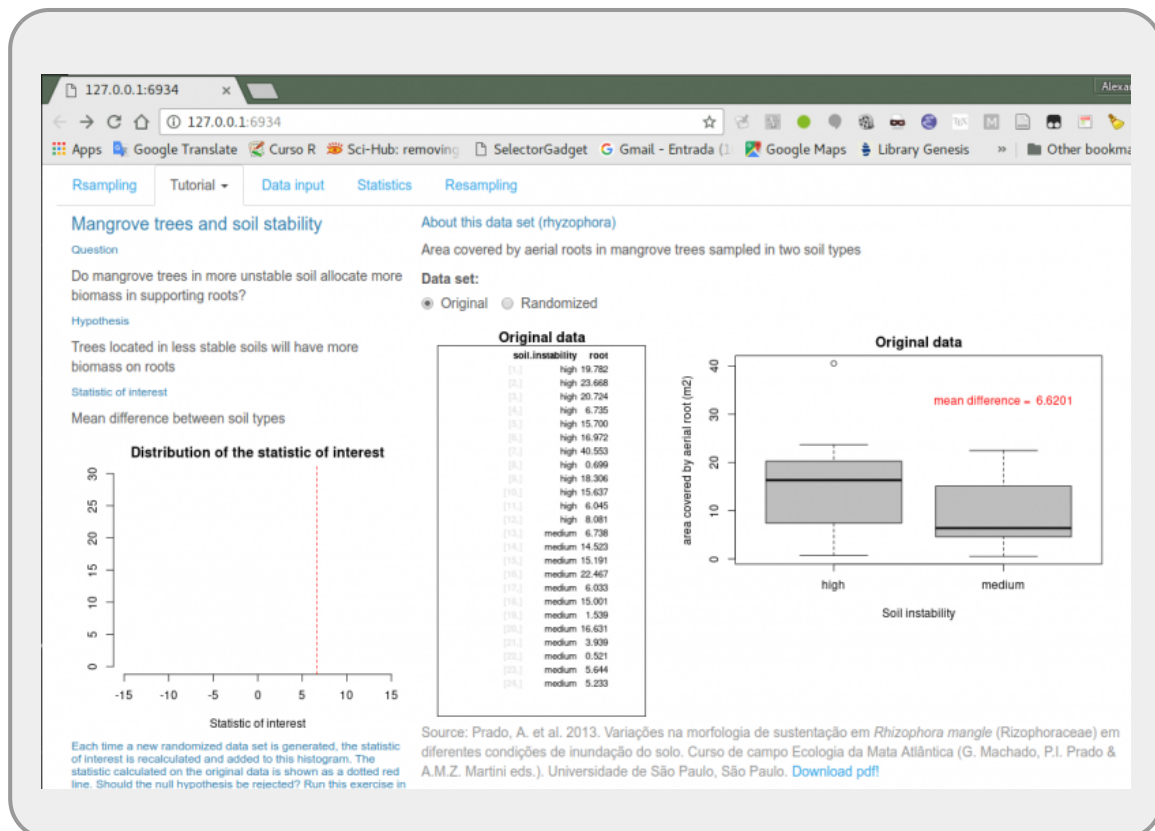
Se tudo correu bem, a seguinte interface deverá estar aberta no navegador:



A interface Rsampling tem 5 abas (Rsampling, Tutorial, Data input, Statistics, Resampling). Explore as abas.

Tutorial Árvores do Mangue

Nesta aula vamos nos ater ao tutorial **“Mangrove trees and soil stability”** na aba *Tutorial*.



- não aperte nenhum botão ainda;
- leia a página do tutorial;
- identifique e tente entender a pergunta e a hipótese colocada;
- identifique qual a hipótese nula para a questão colocada;
- formule uma predição esperada, associada à estatística de interesse;
- entenda o que está expresso na tabela central e no gráfico (boxplot) à direita;
- aguarde a discussão desses tópicos com toda a equipe.

Coleta de dados

A descrição da coleta de dados do trabalho original é a seguinte:

“Desenvolvemos o estudo em uma área de manguezal às margens do Rio Una, no município de Peruíbe, litoral sul do estado de São Paulo. Delimitamos duas parcelas de 50 x 50 m, de mesmo relevo, uma próxima à margem do rio, sujeita a um maior tempo de inundação, a qual nos referimos como mais lodosa, e outra no interior do mangue, sujeita a um menor tempo de inundação, a qual nos referimos como menos lodosa. Fizemos uma amostragem arbitrária de 12 indivíduos de *Rhizophora mangle* em cada uma das parcelas. As variáveis utilizadas como medidas de sustentação dos indivíduos foram área de fixação da raiz e número de pontos de apoio de raízes escora no solo. Para medir a área de fixação das raízes de cada indivíduo, aproximamos a área abrangida pelos pontos de apoio das raízes no solo ao formato de uma elipse. Medimos, em metros, o raio maior e o menor dessa elipse. O número de pontos de apoios foi obtido contando todas as vezes que as ramificações das raízes tocavam no solo....”

Definindo um cenário nulo

Entender a hipótese e sua conexão com a estatística de interesse é crucial para o entendimento do que segue. Faça uma breve pausa e verifique com colegas da turma se entenderam o significado desses conceitos. Iremos resgatar esses conceitos em outros momentos durante o curso, mas não deixe a aula sem os ter entendido.

Agora precisamos definir um cenário associado à hipótese nula, o nosso cenário nulo. Imagine que pudéssemos repetir a amostragem, seguindo o mesmo protocolo, muitas vezes. Quais valores da estatística de interesse seriam mais frequentes partindo da premissa que nosso cenário nulo é verdadeiro?

Será que podemos simular o cenário nulo com nossos próprios dados?

- na janela do tutorial selecione “Randomize”, depois volte para “Original”. Alterne entre essas opções algumas vezes e procure entender o que acontece quando muda a seleção;
- deixe “Randomize” selecionado e clique no botão “Do it again” algumas vezes;
- compare os resultados dos valores que aparecem no boxplot com os valores que são colocados no gráfico à esquerda;
- clique em “Do it again” várias vezes, até que umas das barras chegue ao valor máximo do gráfico à esquerda;
- discuta com os colegas o significado desse gráfico e aguarde a discussão com toda a equipe.

Reamostrando os dados

Caso tenha entendido o tutorial, podemos partir para o teste de hipótese por aleatorização.

- Selecione os dados de rizophora na aba *Data input*;
- rode o teste passando pela seleção da estatística de interesse e depois pelo processo de reamostragem
- faça as escolhas das opções discutindo com os colegas o significado de cada uma;
- quando estiver seguro das opções rode pelo menos 2500 reamostragens (trials) e interprete o resultado;
- mude as opções na barra de rolagem **Alternative**;
- discorra com os colegas o significado de cada uma das opções do **Alternative**;
- aguarde para discutir com os professores o significado dessas opções.

1)

se não sabe a diferença entre os dois tipos de arquivos, provavelmente irá preferir o zip

From:

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

Permanent link:

<http://labtrop.ib.usp.br/doku.php?id=cursos:planeco2017:roteiro:02-hipot>



Last update: **2018/03/05 12:12**