

TRATAMENTO DE DADOS CENSURADOS EM ESTUDOS AMBIENTAIS

Cristiano Christofaro^{a,*} e Mônica M. D. Leão^b

^aDepartamento de Engenharia Florestal, Faculdade de Ciências Agrárias, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Rodovia MGT 367 - Km 583, 5000 Diamantina – MG, Brasil

^bDepartamento de Engenharia Sanitária e Ambiental, Escola de Engenharia, Universidade Federal de Minas Gerais, Av. Antônio Carlos, 6627, Belo Horizonte – MG, Brasil

Análises de dados censurados no pacote NADA (Nondetects And Data Analysis)

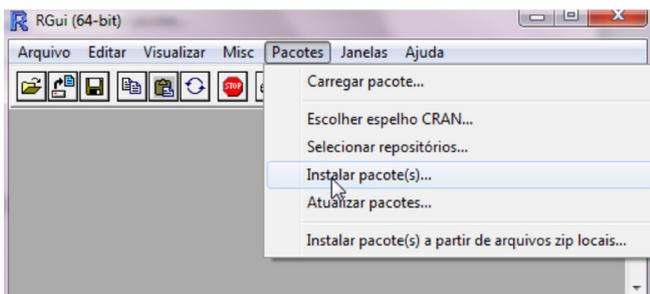
R é uma linguagem e um ambiente utilizado para análises estatísticas e na produção de gráficos. O usuário pode realizar estatísticas a partir de funções existentes (já incluídas no pacote básico ou via pacotes extras) ou pela programação de novas rotinas/extensões. Para a análise de dados abaixo do limite de detecção, o pacote mais conhecido e utilizado é denominado NADA (*Nondetects And Data Analysis for environmental data*)¹ que realiza análises estatísticas mais comumente aplicadas em pesquisas nesse campo.

Abaixo os principais passos necessários para rodar o pacote NADA no programa R pela primeira vez em ambiente Windows:

1 – Vá ao site do R e instale a versão mais atual do programa, conforme instruções no site <http://www.r-project.org/>

2 – Instalação do pacote NADA (*Nondetects And Data Analysis for environmental data*): O programa R já vem com pacotes básicos. Análises mais especializadas requirem pacotes especiais. Esses pacotes precisam ser instalados apenas uma vez. Para realizar essa instalação vá em:

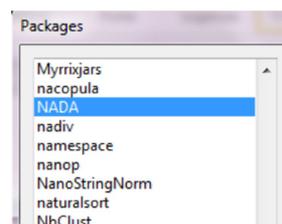
Pacotes -> Instalar Pacote(s)...



Selecione agora sua localização (ou qualquer outro local) e clique OK.



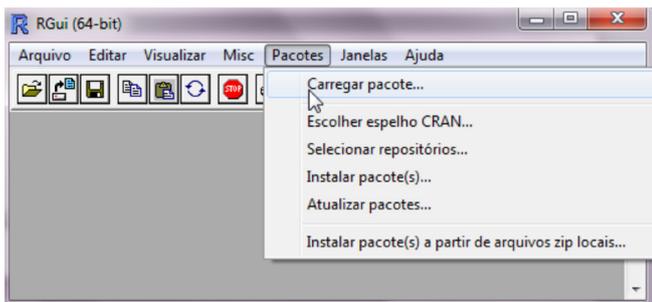
Selecione na lista o pacote para instalar, nesse caso NADA, e clique OK.



*e-mail: cristiano.christofaro@ufvjm.edu.br

5. Carregue o pacote NADA, necessário para os dados censurados:

Pacotes -> Carregar Pacotes...



Selecione na lista NADA e clique OK.

Outra forma de carregar o pacote NADA é escrevendo o comando abaixo no R console:

```
> library(NADA)
```

Obs: O carregamento do pacote, por qualquer um dos dois métodos, deve ser feito toda vez que o R for aberto.

6. Leitura dos dados. Para uso no pacote NADA os dados precisam apresentar uma formatação específica: devem haver duas colunas para cada variável, uma com os resultados das amostras e outra com o código TRUE para indicar os valores censurados e FALSE para os não censurados. No exemplo abaixo, a coluna 'Pb' apresenta as concentrações de chumbo na água e a coluna 'D_Pb' indica presença (TRUE) ou ausência (FALSE) de censura para a respectiva amostra.

Pb	D_Pb
0.0370	FALSE
0.0050	TRUE
0.1350	FALSE
0.0080	FALSE
0.0210	FALSE
0.0090	FALSE
0.0090	FALSE
0.0330	FALSE
0.0050	TRUE
0.0180	FALSE
0.0150	FALSE
0.0120	FALSE
0.0130	FALSE
0.0050	TRUE
0.0050	TRUE
0.0120	FALSE
0.0200	FALSE
0.0050	TRUE

Após gravar os dados no formato 'txt' o comando abaixo deve ser utilizado para a leitura dos dados pelo R:

```
dados.Pb = read.table("C:/Chumbo.txt", header = TRUE)
```

Obs:

- Nesse exemplo, os dados foram originalmente salvos com o nome "Chumbo.txt" e foram gravados no endereço 'C:/'. O objeto 'dados.Pb' passa a ser a referência do R para essas variáveis. O usuário pode escolher qualquer nome para esse objeto e deve alterar o endereço do arquivo de acordo com a pasta onde o arquivo '.txt' foi salvo.

- O programa diferencia letras maiúsculas e minúsculas. Então para o R "Chumbo.txt" é diferente de "chumbo.txt" e "dados.Pb" é diferente de "Dados.Pb".

7. A partir de então é possível rodar as análises disponibilizadas no pacote NADA. Os nomes após o caracter '\$' correspondem à coluna de interesse no objeto com os dados (nesse caso, dados.Pb). Alguns comandos importantes desse pacote são:

```
# criar boxplot
> cenboxplot(dados.Pb$Pb, dados.Pb$D_Pb)
# cálculo do mínimo, máximo, número de dados censurados e percentual de censura
> censummary(dados.Pb$Pb, dados.Pb$D_Pb)
# cálculo da mediana, média e desvio padrão (por KM, ROS e MLE)
> censtats(dados.Pb$Pb, dados.Pb$D_Pb)
```

Outras opções de configuração e questões para entendimento dos comandos acima podem ser visualizadas utilizando o comando help('nome do comando').

Exemplos: help(cenboxplot), help(censtats)

8 - Outras informações e textos sobre o R podem ser visualizadas em documentos *online* disponíveis site oficial: www.r-project.org e <http://cran.r-project.org/web/packages/NADA/index.html>

REFERÊNCIAS

1. Lee, L.; *NADA: Nondetects And Data Analysis for Environmental Data*, R package version 1.5-5. <http://CRAN.R-project.org/package=NADA>, 2013