

ESTATÍSTICA APLICADA ÀS CIÊNCIAS AMBIENTAIS

Prof. Anderson Rodrigo da Silva

<< Exercícios em R >>

#1 – Anderson’s Iris Data

Medidas (cm) das variáveis comprimento e largura de sépala, comprimento e largura de pétala de 50 flores de cada uma das três espécies do gênero *Iris* (lírio): *Iris setosa*, *I. versicolor* e *I. virginica*. O objetivo é estudar a variabilidade morfológica destas espécies. Dados disponíveis no R. Digite `data(iris)` ou `help(iris)`

- Crie um objeto chamado “`versi`” para armazenar apenas os dados da espécie *versicolor*, um subconjunto de dados. Tente fazer isto usando a função `subset`. Na dúvida, tente observar os exemplos executando `help(subset)`
- Construa box-plots para verificar se há outliers em cada uma das variáveis (colunas) de “`versi`”
- Calcule medidas de posição (mean, median) e de dispersão (var, sd) para cada variável de “`versi`”
- Calcule o coeficiente de variação para cada variável de “`versi`”
- Use a função `aggregate` para automatizar o cálculo da média de todas as variáveis de iris por espécie. Dica: `help(aggregate)`
- Construa um histograma (`hist`) ou gráfico de densidade kernel (`density`) da variável `Sepal.Length` (comprimento de sépala) de “`versi`” e salve em formato `.png`. Verifique pelo teste de Shapiro-Wilk se essa variável segue distribuição normal. Dica: `help(shapiro.test)`
- Avalie a hipótese de que as variâncias das espécies *setosa* e *versicolor* são estatisticamente ($\alpha = 0,05$) iguais. Dica: `help(var.test)`
- Aplique o teste t (`t.test`) adequado (variâncias iguais ou diferentes) para comparar as médias das espécies *setosa* e *versicolor* em termos de `Sepal.Length`
- Responda a questão anterior construindo intervalos de 90% de confiança para a média de `Sepal.Length` cada espécie. Para tal, execute o teste t para uma média (uma espécie por vez).

#2 – Teste t para amostras pareadas

Doze cobaias adultas foram submetidas a certa dieta durante uma semana. Os animais foram identificados, sendo mantidos em gaiolas individuais. Os pesos, em gramas, no início e no fim da semana são dados a seguir. Armazene os dados de cada amostra em objetos com nome “`inicio`” e “`fim`” e, em seguida, execute o teste t (`t.test`) apropriado para verificar efeito ($\alpha = 0,05$) da dieta sobre o peso.

Início	635	703	660	550	602	735	678	565	623	629	620	725
Fim	640	711	673	548	609	730	687	575	625	642	618	735

#3 – Importação de dados

O R permite importar arquivos em diversos formatos, dentre eles: `.R`, `.txt`, `.dat`, `.csv` (Excel). Para tanto, há algumas principais funções. Ei-las:

- `read.table()` - para leitura de dados tabulados em colunas
- `read.csv()` ou `read.csv2()` - para leitura de dados tabulados em colunas no formato csv

Transfira os dados da questão anterior para uma planilha do Excel, tabulando-os em colunas (Início, Fim), e salve em formato `.csv` (separados por vírgula). Copie os dados do Excel e cole em um arquivo do Bloco de Notas (formato `.txt`). Em seguida, use as funções mencionadas para importar esses arquivos para o R. Crie um objeto chamado “`cobaias`”.