

Estatística Básica

Introdução à Análise Exploratória de Dados

Renato Dourado Maia

Instituto de Ciências Agrárias

Universidade Federal de Minas Gerais



Ferramentas Computacionais

- ❑ Foi comentado, no início do curso, que o desenvolvimento de ferramentas computacionais foi de extrema importância para a difusão e utilização de métodos estatísticos.
- ❑ Muitos programas de uso comum, como as planilhas eletrônicas, contêm rotinas estatísticas incorporadas às suas funções básicas.
- ❑ Existem diversos pacotes estatísticos, que são programas especificamente desenvolvidos para efetuar análises estatísticas.

Ferramentas Computacionais

- **Independente** do programa, a sua utilização pode ser resumida em **três etapas**:
 - Entrada de Dados.
 - Execução da Análise Estatística.
 - Interpretação de Resultados.

Ferramentas Computacionais

- A **Entrada de Dados** deve assumir certas **convenções**, que podem **variar** de um programa para outro, e tem como efeito a criação de uma **matriz**, em que, normalmente, cada **linha** corresponde a uma **unidade experimental** (elemento da população ou amostra), e cada **coluna** a uma **variável**.
- Para o exemplo estudado na aula passada, os dados seriam representados por uma matriz de dimensão 50 (linhas) por 14 (colunas).

Ferramentas Computacionais

- ❑ A Execução da Análise Estatística pressupõe o conhecimento de como o programa que está sendo utilizado trabalha as informações: as funcionalidades disponíveis, a linguagem de programação, as convenções, etc...
- ❑ Após a execução da análise estatística, chega-se à etapa de Interpretação de Resultados. O detalhe importante é: o computador não interpreta os resultados! 😊

Ferramentas Computacionais

- Os *softwares* de planilhas eletrônicas possuem funções para efetuar análises estatísticas, sendo muito utilizados em função de sua interface ser visual e intuitiva, especialmente para a utilização das ferramentas mais básicas.
 - Dentre eles, o mais conhecido é o Microsoft Excel, que faz parte do pacote Microsoft Office, uma suíte de aplicativos proprietária.
 - Uma alternativa livre é o Calc, que faz parte da suíte de aplicativos LibreOffice, similar ao Microsoft Office.
 - Site de referência (acessado em 07/08/2013):
 - <http://pt-br.libreoffice.org/>

Ferramentas Computacionais

- É possível fazer com que as **planilhas eletrônicas** possam ser mais poderosas para análises estatísticas.
- Na **página da disciplina** há *links* para o conjunto de macros **Ooo Statistics**, que agrega funcionalidades para o **Calc**, e para o **Action**, um *software* que permite que, por intermédio do **Excel**, possam ser realizadas análises estatísticas no **R** (falaremos daqui a pouco sobre o R).
 - No *site* do **Action** há um material interessante sobre estatística! 😊

Ferramentas Computacionais:



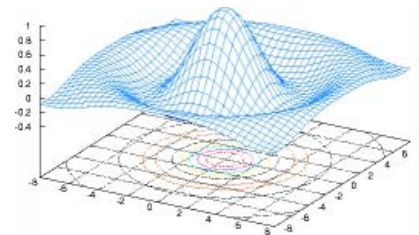
- ❑ Um programa largamente utilizado para Computação Científica é o Matlab.
- ❑ O Matlab possui um pacote (*Statistics Toolbox*) para efetuar análises estatísticas.
- ❑ O programa é excelente, mas é proprietário...
- ❑ Site de referência (acessado em 07/08/2013):
 - <http://www.mathworks.com/>

Ferramentas Computacionais: Scilab



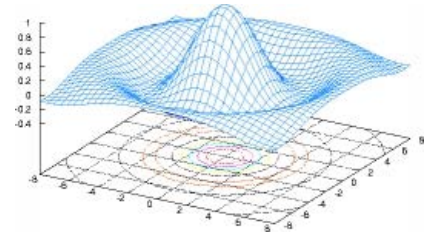
- ❑ O **Scilab** também possui um pacote para efetuar análises estatísticas.
- ❑ É um **excelente** programa para **Computação Científica**, e que vem sendo cada vez mais utilizado, pois, ao contrário do **Matlab**, é **livre**.
- ❑ *Site* de referência (acessado em 07/08/2013):
 - <http://www.scilab.org/>

Ferramentas Computacionais:



- ❑ O **Octave** é um “clone” gratuito do **Matlab**, e também possui um pacote para efetuar análises estatísticas.
- ❑ É interessante ressaltar que o **Octave busca ser compatível com o Matlab**, de modo que as funções disponíveis tenham a mesma interface, e as linguagens de programação sejam similares.
- ❑ O Octave é fornecido para ser utilizado num **terminal/console**, mas **há interfaces gráficas disponíveis**, dentre as quais destaco o **GUIOctave**.

Ferramentas Computacionais:



- *Sites* de referência (acessados em 07/08/2012):
 - <http://www.gnu.org/software/octave/>
 - <https://sites.google.com/site/gui octave/>

Ferramentas Computacionais:



- ❑ O **PSPP** é um *software* para análises estatísticas que pode ser considerado um **substituto livre** para o **SPSS**, um *software* proprietário: com algumas poucas exceções, a utilização do PSPP é igual à do SPSS, sendo que os **arquivos de dados** são **compatíveis**...

- ❑ *Site* de referência (acessado em 07/08/2013):
 - <http://pspp.michelboaventura.com/>

Ferramentas Computacionais:



- ❑ O R é um *software* para **Computação Científica**, muito utilizado para **análise estatística**, desenvolvido na Universidade de Auckland, na Nova Zelândia.
- ❑ Ele tem sido **extensivamente utilizado** pela **comunidade científica** devido à sua versatilidade, e por ser **livre**.
- ❑ *Site* de referência (acessado em 07/08/2013):
 - <http://www.r-project.org/>

Ferramentas Computacionais:



- ❑ Na página da disciplina, disponibilizarei *links* e materiais para **apoiar** o aprendizado sobre o programa.

- ❑ O próprio *site* de referência disponibiliza **muitos materiais**, mas destaco que a **ajuda do programa** é muito **boa** e **prática**. Com a instalação completa, a ajuda fornece:
 - FAQ do R e FAQ do R para Windows.
 - Manuais em PDF.
 - Ajuda HTML.
 - Etc...

Ferramentas Computacionais

- Vamos ver **rapidamente** a interface de algumas das ferramentas computacionais apresentadas...
- As figuras apresentadas na aula passada foram geradas pelo **Matlab** e pelo **Octave**. Caso o tempo permita, veremos os códigos que escrevi e rodaremos cada um deles...

Ferramentas Computacionais

- Agora, vamos brincar um pouco com o R:
 - Carregar os dados do exemplo (Questionario.txt).
 - Construir um *box-plot* para a variável Peso.
 - Construir um histograma para a variável Peso.
 - E o que mais conseguirmos no tempo que restar! 😊

Na medida do possível, comentarei sobre as funcionalidades do R e sobre como utilizá-las... Mas para aprender mesmo a utilizar o programa, cada um deverá explorar a ajuda: lembrem-se da necessidade de desenvolvimento de um perfil mais científico, dando atenção especial à construção de uma boa capacidade autodidata e de análise, enfim, um perfil de pesquisador...