

BIOESTATÍSTICA

AULA 1

Anderson Castro Soares de Oliveira
José Nilton da Cruz

Departamento de Estatística/ICET/UFMT

Apresentação do Curso

APRESENTAÇÃO DO CURSO

- Os conteúdos estarão disponíveis no site www.andersoncso.webnode.com.br
- Serão realizadas 3 avaliações: duas avaliações escritas e um seminário, sendo:

Avaliação	Peso
Avaliação Escrita 1	0,35
Avaliação Escrita 2	0,35
Seminário	0,30

PROGRAMA DAS AULAS

- Introdução a estatística - Noções básicas: população, amostra, tipos de variáveis
- Introdução ao planejamento de pesquisa: hipótese, planejamento, observações, análise e conclusões
- Construção de tabelas e gráficos
- Medidas de Posição e Dispersão
- Noções de Probabilidade
- Conceito de variáveis aleatórias discreta e contínua
- Distribuições de probabilidade: Binomial, Normal, t-student, Qui-quadrado e F-Snedecor

PROGRAMA DAS AULAS

- Amostragem: plano amostral
- Principais tipos de amostragem: Amostragem Aleatória Simples, Amostragem Estratificada, Amostragem por Conglomerados
- Cálculo do tamanho de amostra
- Princípios de Estimação: Estatísticas e Parâmetro
- Intervalos de confiança para a média, proporção, diferença entre duas médias e diferença entre duas proporções
- Teste de hipótese: Procedimento Geral do Teste de Hipóteses; Testes para a média; Testes para a proporção.

PROGRAMA DAS AULAS

- Regressão: definição; estimativa de parâmetros; fontes de variação na Análise de Regressão; critérios para a seleção de modelos.
- Medidas de associação: teste de hipóteses de associação
- Tipos de estudos epidemiológicos: Estudo de prevalência, Estudo caso-controle e Estudo de coorte
- Medidas de efeito (Medidas de risco): Razão de Prevalência, Razão de chance (odds ratio) e Risco Relativo.

BIBLIOGRAFIAS

- BERQUÓ E.S.; SOUZA J.M.P.; GOTLIEB S.L.D. Bioestatística. EPU. São Paulo, 1984.
- CALLEGARI-JACQUES, S. M. Bioestatística: princípios e aplicações. Porto Alegre: Artmed, c2003. 255 p.
- MOTTA, V. T. Bioestatística. 2. ed. Caxias do Sul: EDUCS, 2006. 190 p.
- PAGANO, M.; GAUVREAU, K. Princípios de Bioestatística. 1. ed. São Paulo: Cengage Learning, 2004. xv, 506 p.
- PETRIE, A. Estatística em ciência animal e veterinária. São Paulo: Ed. Roca. 2009, 248p.
- SAMPAIO, I.B.M. Estatística aplicada à experimentação animal. 3a. ed. Belo Horizonte: Editora da UFMG. 2007, 265p.

SITES IMPORTANTES

- R - www.cran.r-project.org.

Introdução

O QUE É ESTATÍSTICA

- A palavra estatística vem do latim *STATUS* - que significa estado.
- Estatística é um conjunto de conceitos e métodos científicos para coleta, organização, descrição, análise e interpretação de dados experimentais, que permitem conclusões válidas e tomadas de decisões razoáveis.

UTILIZAÇÃO DA ESTATÍSTICA

- A utilização da Estatística é cada vez mais acentuada em qualquer atividade profissional.
- Nos seus diversificados ramos de atuação, as pessoas estão frequentemente expostas à Estatística, utilizando-a com maior ou menor intensidade.
- Pak & OH (2010) ¹ fizeram um levantamento de 129 do Journal of Veterinary Clinics no período de 2006 a 2007.
 - 94 dos artigos tinham estatística
 - 66 dos artigos tinham erros na estatística utilizada

¹Pak & OH (2010). Statistical Issues in the Articles Published in the Journal of Veterinary Clinics. Journal of Veterinary Clinics 27(2) : 170-174

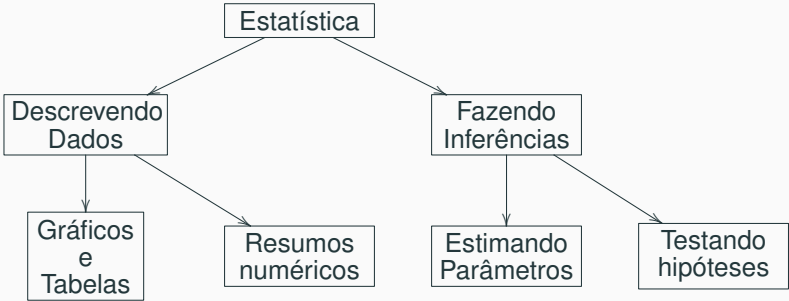
ABUSOS DA ESTATÍSTICA

- São muitos os abusos da estatística, quando os dados são apresentados de forma enganosa.
- Há três tipos de mentira: as mentiras, as mentiras sérias e a estatística. (Benjamin Disraeli)
- Algumas pessoas usam a estatística “como um bêbado utiliza um poste de iluminação - para servir de apoio e não para iluminar” (Andrew Lang)
- “É fácil mentir com estatística, difícil é dizer a verdade sem ela” (Darrell Huff no livro Como Mentir com Estatística)

RAMIFICAÇÕES DA ESTATÍSTICA

- Estatística descritiva - tem por objetivo organizar, apresentar e sintetizar dados observados de determinada população por meio de tabelas, gráficos e medidas descritivas, sem pretensões de tirar conclusões de caráter extensivo.
- Estatística Inferencial - Métodos que tornam possível a estimação de características de uma população baseada nos resultados amostrais no intuito de tirar conclusões.

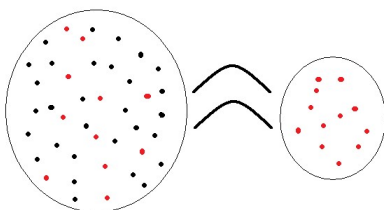
RAMIFICAÇÕES DA ESTATÍSTICA



Conceitos Básicos

CONCEITOS BÁSICOS

- População (N) - Conjunto de todos os elementos relativos a um determinado fenômeno que possuem pelo menos uma característica em comum, a população é o conjunto Universo.
- Amostra (n) - um subconjunto da população, a amostra deve ser selecionada seguindo regras e deve ser representativa, de modo que ela represente todas as características da população como se fosse uma fotografia desta.



CONCEITOS BÁSICOS

- Censo - atividade de inspecionar (observar) todos os elementos de uma população, com o objetivo de conhecer, com certeza suas características;
- Amostragem - É o processo de formação de informações dos “n” elementos amostrais, no qual deve seguir um método criterioso e adequado (tipos de amostragem).

CONCEITOS BÁSICOS

- Variável: é toda característica que, quando observada, pode variar de um indivíduo para outro e podendo assumir valores numéricos ou não-numéricos.
- Definir bem as variáveis de interesse é muito importante, pois essas vão determinar o tipo de análise estatística.
- Pode-se classificar as variáveis para o estudo estatístico em Qualitativas e Quantitativas

VARIÁVEIS QUALITATIVAS

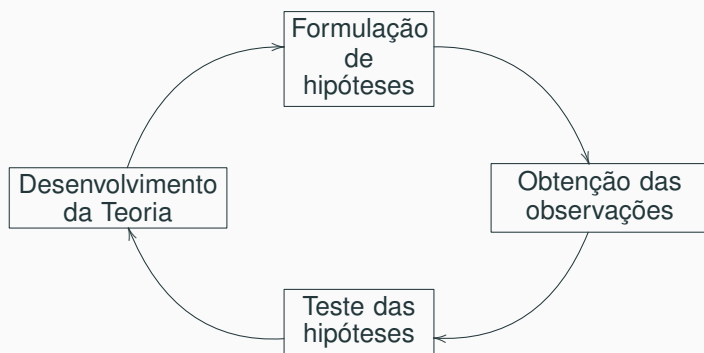
- Variáveis Qualitativas (ou categóricas) - são as características que não possuem valores quantitativos, mas, ao contrário, são definidas por várias categorias, ou seja, representam uma classificação dos indivíduos
 - Variáveis nominais: não existe ordenação dentre as categorias.
Exemplos: Sexo, Cor, Raça, Procedência, doente/sadio, etc.
 - Variáveis ordinais: existe uma ordenação entre as categorias.
Exemplos: Classificação quanto ao porte (grande, médio, pequeno), estágio da doença (inicial, intermediário, terminal), mês de observação (janeiro, fevereiro,..., dezembro).

VARIÁVEIS QUANTITATIVAS

- Variáveis Quantitativas - são variáveis que podem ser medidas em uma escala quantitativa, ou seja, apresentam valores numéricos
 - Variáveis discretas: são aquelas variáveis que podem assumir somente valores inteiros num conjunto de valores. É gerada pelo processo de contagem
Exemplos: número de filhos, números de placas com bactérias, número de análises realizadas pelo laboratório
 - Variáveis contínuas: são aquelas variáveis que podem assumir um valor dentro de um intervalo de valores. É gerada pelo processo de medição
Exemplos: pressão arterial, peso de animais; volume de gás produzido por um biodigestor; etc

Pesquisa Científica

MÉTODOS CIENTÍFICOS



MÉTODO CIENTÍFICO

- Todo estudo estatístico pode ser decomposto em pelo menos duas fases:
 - Reunião ou coleta dos dados estatísticos
 - Análise e interpretação de dados.
- A coleta de dados pode ser realizada de duas formas
 - estudo observacional - por simples observação dos fenômenos em que se tem interesse;
 - estudo experimental - por experimentação, ou seja, provocando-se voluntariamente a aparição de certos fenômenos controlados

MÉTODO CIENTÍFICO

- As pessoas tentam solucionar uma vasta gama de problemas utilizando o pensamento estatístico.
 - Agricultura (qual variedade cresce melhor?)
 - Economia (como estão mudando os padrões de vida?)
 - Pesca (as cotas atuais de peixe, impostas à indústria pesqueira são apropriadas?)
 - Psicologia (como estão relacionadas a timidez e a solidão?)
 - Meteorologia (o aquecimento global é uma realidade?)

PLANEJAMENTO DA PESQUISA

- O que é uma pesquisa?
 - Processo de investigação científica com o objetivo de comprovar ou contestar uma teoria;
 - Uma forma do Pesquisador tentar compreender um fenômeno, por meio de um processo lógico de investigação
- O planejamento e a execução de uma pesquisa fazem parte de um processo sistematizado que compreende várias etapas

PLANEJAMENTO DA PESQUISA

Etapas do planejamento de uma pesquisa

- 1) Escolha do tema;
 - Qual o meu interesse?
 - O que vou pesquisar?
- 2) Revisão de literatura;
 - O que já foi publicado sobre o tema?
- 3) Justificativa;
 - Por que estudar o tema?
- 4) Formulação do problema/hipóteses;
 - O que resta de dúvida sobre o tema?
 - O que ainda está em aberto?
- 5) Determinação de objetivos;
 - O que pretendo alcançar?

Etapas do planejamento de uma pesquisa

6) Metodologia;

- Como se procederá a pesquisa?
- Quais os métodos para obtenção dos resultados?

7) Resultados e Impactos Esperados;

- Quais as possíveis implicações?
- Como será a divulgação e a disseminação do conhecimento (artigo científico, dissertação ou tese)?

Pesquisa Quantitativa × Qualitativa

- Uma vez definido o tema da pesquisa, deve-se escolher entre realizar uma pesquisa qualitativa ou uma quantitativa.
- A pesquisa qualitativa é caracterizada por ser exploratória, e é utilizada quando se busca percepções e entendimento sobre a natureza geral de uma questão, abrindo espaço para a interpretação. Ela não depende de análise estatística para suas inferências.
- Tipos de pesquisa qualitativa
 - Estudo de Caso
 - Observação participante

Pesquisa Quantitativa × Qualitativa

- A pesquisa quantitativa é um estudo estatístico que se destina a descrever as características de uma determinada situação, medindo numericamente as hipóteses levantadas a respeito de um problema de pesquisa.
- O pesquisador deve possuir um conhecimento razoável de estatística para desenvolver uma pesquisa quantitativa ou, então, consultar um estatístico para auxiliá-lo

PESQUISA ESTATÍSTICA

- Uma pesquisa estatística consiste em um trabalho de identificação, reunião, tratamento, análise e apresentação de informações (dados) para atender uma certa demanda
- A realização de uma pesquisa estatística envolve a coleta e o resumo de informações das características de interesse de unidades de uma população, com o objetivo de estimar características da população e testar hipóteses.

- No planejamento da pesquisa estatística deve-se definir:
 - População alvo
 - Variáveis e a forma de coleta de dados
 - Métodos para resumo e análise dos dados.

Definição da População de Pesquisa

- Uma parte importante do delineamento de qualquer pesquisa é a definição da população
- A definição da população dependerá obviamente dos objetivos da pesquisa, das características a mensurar, dos recursos disponíveis.
- População alvo: população para a qual se deseja obter informação
- População de pesquisa: população realmente coberta pela pesquisa
 - Pesquisas: Estatísticas do Febre Aftosa em Mato Grosso
 - População alvo: Bovinos das propriedades rurais do Estado de Mato Grosso.
 - População de pesquisa: Bovinos das propriedades rurais do Estado de Mato Grosso que participam do programa de vacinação.

Definição das Variáveis

- Ao estudar um determinado fenômeno é necessário determinar as características que serão analisadas: as variáveis
- As variáveis são características que podem ser observadas ou medidas em cada elemento pesquisado.
- As hipóteses serão formuladas no estudo a partir da definição das variáveis
- A mesma variável pode ser mensurada de diversas formas, levando a diferentes conclusões, assim deve-se definir:
 - nível de mensuração - o quanto de informação cada variável apresenta e se ela será qualitativa ou quantitativa
 - nível de manipulação - como uma variável relaciona e com as outras no estudo

Definição da Coleta dos Dados

- Há duas formas básicas de coletar os dados: por censo ou amostragem
- Além da decisão por censo ou amostragem deve-se decidir se será utilizado dados primários ou secundários.
 - dados primários: são aqueles coletados pela primeira vez pelo pesquisador;
 - dados secundários: são dados existentes, coletados por outros pesquisadores e disponíveis em relatórios ou publicações.
- Outro ponto importante é a definição dos instrumentos de coleta de dados

Definição da Coleta dos Dados

- Os instrumentos de coleta de dados são as ferramentas que permitirão a coleta, o levantamento de dados, a produção de informações.
- Os métodos de obtenção das informações são os mais diversos e dependem de vários fatores, sendo os principais o objeto de estudo, o fenômeno a ser estudado e o alcance que se pretende
- Os instrumentos podem ser questionários, experimentos ou observações.

Questionário

- O questionário constitui-se num instrumento muito utilizado no processo de obtenção/coleta de informações.
- A elaboração de um questionário baseia-se em experiências anteriores e no bom senso:
- É importante a realização de pré-teste do questionário e demais instrumentos de coleta;
- Ao aplicar um questionário pode-se cometer dois tipos de erros:
 - Devido ao questionário: Tamanho, ordem da perguntas, redação das perguntas, utilização de conceitos incorretos ou desatualizados.
 - Devido ao entrevistador: Características do entrevistador, ansiedade, registro incorreto das respostas;

Tipos de delineamento de pesquisa

- Existem várias formas de se delinear uma pesquisa, estas diferem na forma como se obtém os dados.
- Podem-se definir dois tipos de estudos: experimentais e observacionais.
- Em um estudo observacional, observamos e medimos características específicas, mas não tentamos modificar os sujeitos objetos do estudo.
- Em um experimento, aplicamos algum tratamento e passamos, então, a observar seu efeito sobre os sujeitos. (Os sujeitos do experimento são chamados unidades experimentais).

Fontes de Dados Estatísticos

- Fontes de dados estatísticos referem-se às obras que divulgam dados quantitativos, numéricos, resultantes de compilações feitas de maneira sistemática, com periodicidade determinada, voltadas a informações comerciais, administrativas, científicas, industriais, entre outras.
- As fontes que divulgam dados estatísticos e podem ser diversos: Anuários Estatísticos, Relatórios Anuais, Bases de Dados, etc.

Fontes de Dados Estatísticos

- Bases de dados:
 - IBGE <http://www.sidra.ibge.gov.br/>
 - Organização Mundial de Sanidade Animal - <http://www.oie.int>
 - Panaftosa - www.panaftosa.org.br

Introdução ao R

INTRODUÇÃO AO R

O que é o R?

R é uma linguagem e ambiente para computação estatística e gráficos, desenvolvido por Robert Gentleman e Ross Ihaka do Departamento de Estatística da Universidade de Auckland em Nova Zelândia



Robert Gentleman



Ross Ihaka

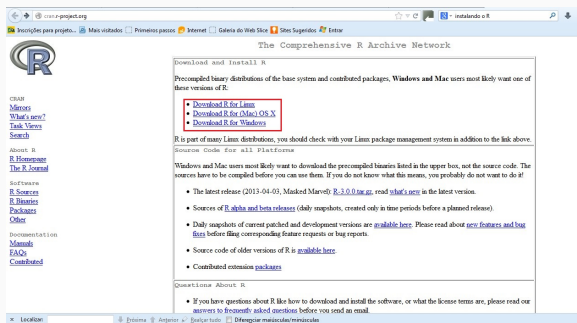
O que é o R?

- Para entender o R é necessário saber que trata-se de uma linguagem;
- O R é compatível com quase todos os sistemas operacionais;
- O R é um programa com pouca interatividade.

INTRODUÇÃO AO R

Instalação

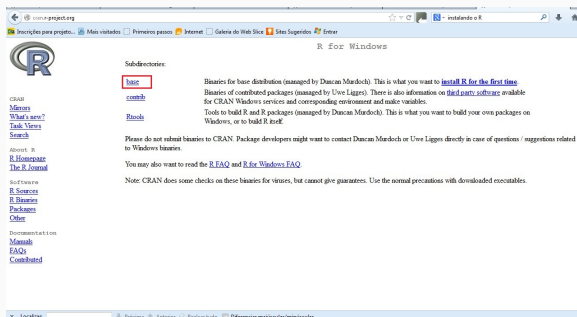
- A instalação do R pode ser feita através do site <http://cran.r-project.org/>.
- Primeiro deve seleccionar o sistema operacional: Linux, Mac ou Windows



INTRODUÇÃO AO R

Instalação

- Para Windows selecione o menu `base` que irá aparecer o link para descarregar o instalador no computador.



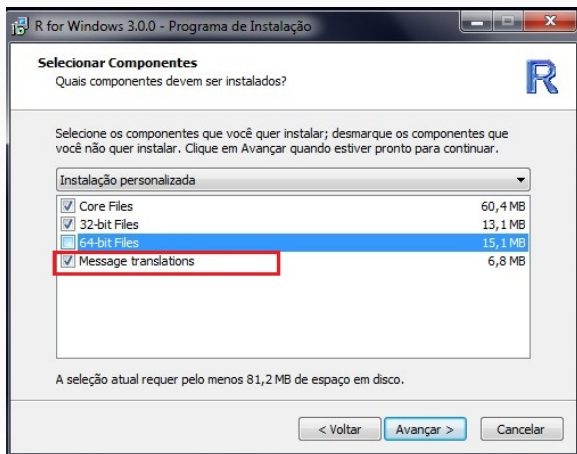
Instalação

- No instalador existem duas versões do R 32 e 64 bits.
- Se o sistema operacional for de 32 bits apenas a primeira versão é instalada
- Se o sistema operacional for de 64 bits as duas versões são instaladas.
- Após o término do download, comece a instalação selecionando o idioma e clique em `Ok`
- Nas demais telas basta escolher a opção `Avançar`

INTRODUÇÃO AO R

Instalação

- Quando chegar na tela **Selecionar componentes** verifique se a opção **Message translation** está marcada.

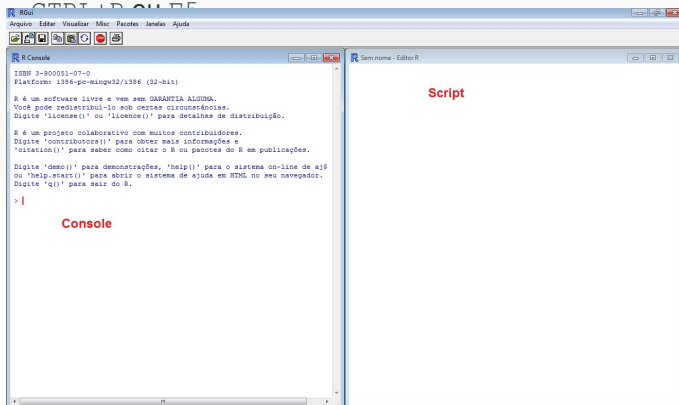


- Nas demais telas basta escolher a opção **Avançar**

INTRODUÇÃO AO R

Usando o R

- No Windows o R apresenta uma interface gráfica com barra de ferramentas no topo;
- Essa interface apresenta o Console e Script.
- Console onde se executam os comandos
- Script é um arquivo, onde você digita todas as análises e comandos
- Para executar os comandos digitados no script, seleciona-se o texto ou linhas de interesse e utiliza-se as teclas



Instalando Pacotes no R

- O R em geral é instalado apenas com as configurações mínimas para seu funcionamento básico (pacote base);
- Para realizar tarefas mais complicadas pode ser necessário instalar pacotes adicionais (packages ou library);
- A lista de todos os pacotes podem ser encontrados no site <http://cran.r-project.org/>.

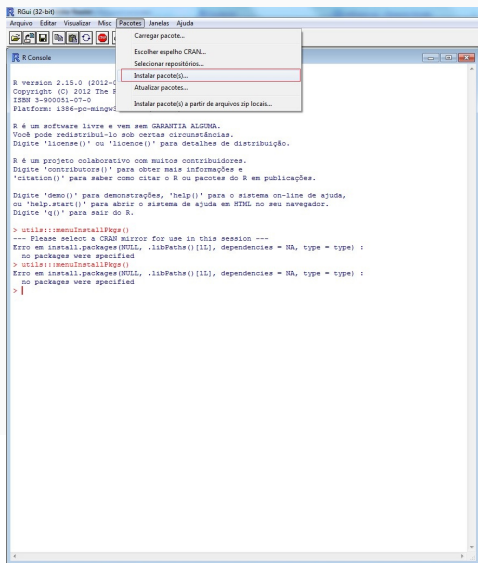
Instalando Pacotes no R

- Para instalar um pacote, deve estar conectado a internet, assim:
 - Vá na barra de ferramentas `Pacotes` em seguida em `Instalar pacote(s)`;
 - Em seguida irá abrir uma janela de espelho do CRAN (base de pacotes).
 - Escolhido o espelho do qual irá baixar o pacote, irá aparecer uma lista de pacotes, assim escolhe qual quer instalar.

INTRODUÇÃO AO R

Instalando Pacotes no R

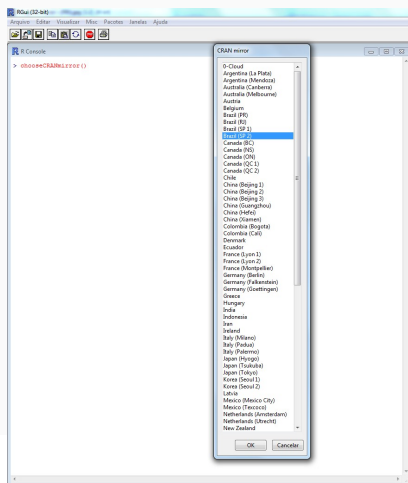
- Vá na barra de ferramentas Pacotes e em seguida em Instalar pacote(s);



INTRODUÇÃO AO R

Instalando Pacotes no R

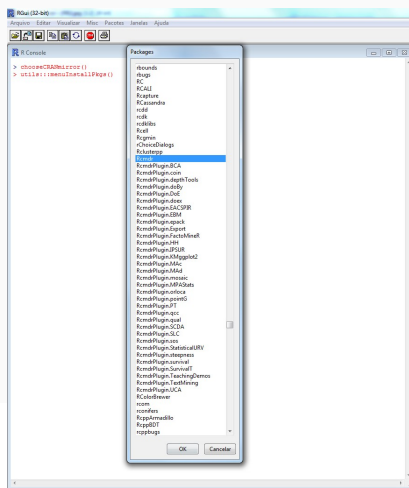
- Em seguida irá abrir uma janela de espelho do CRAN (base de pacotes).



INTRODUÇÃO AO R

Instalando Pacotes no R

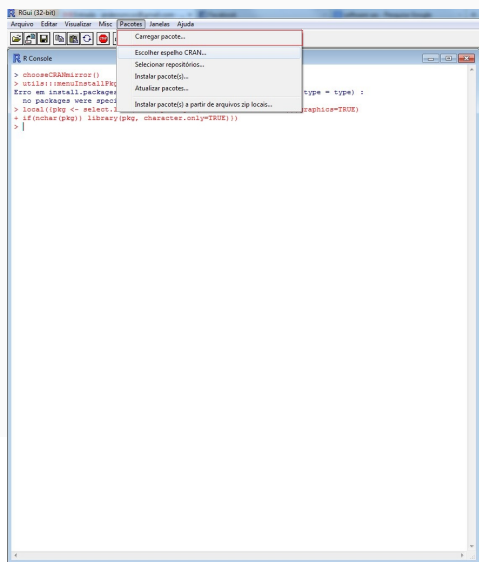
- Escolhido o espelho do qual irá baixar o pacote, irá aparecer uma lista de pacotes, assim escolhe qual se deseja instalar.

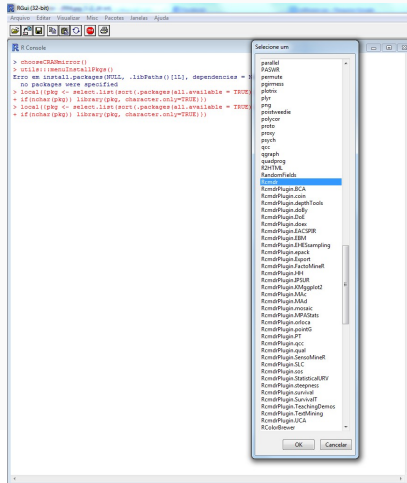


INTRODUÇÃO AO R

Utilizando um Pacote Instalado

- Vá na barra de ferramentas Pacotes e em seguida em Carregar pacote(s);





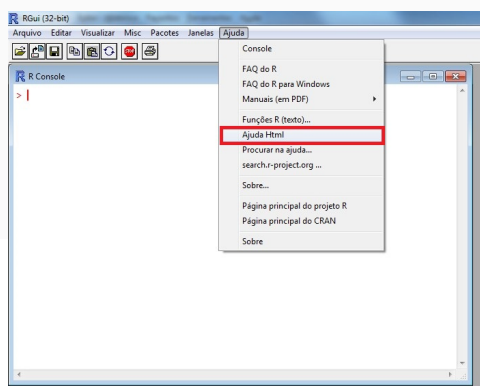
Ajuda e Documentação

- Sistema de ajuda;
- Ajuda online;
- Manuais e publicações eletrônicas trabalho sob a forma de livros, etc.

INTRODUÇÃO AO R

Sistema de ajuda

- O sistema de ajuda é uma coleção de páginas de manual que descreve cada função para o usuário.
- Quando é instalado o R, é criado um diretório de "ajuda", que consta alguns manuais básicos e manuais para os pacotes instalados
- Na barra de ferramentas Ajuda em seguida em Ajuda Html;



Sistema de ajuda

- Para ver a ajuda relativa a uma função basta utilizar os comandos: `help("função")` ?função no lugar de função coloca-se o nome da função que se quer ajuda,
- ajuda sobre a média, utiliza-se `help("mean")`
- Para ver informações relativa a um pacote utiliza-se o comando `help(package = "nome pacote")`

INTRODUÇÃO AO R

Ajuda on line

- A ajuda online possui diversas informações relativos a funções e perguntas relativas a utilização do R.
- Para ver ajuda online deve estar conectado a internet e utilizar o comando:
`RSiteSearch("nome de algo pra ajuda")`
- Para obter ajuda on line sobre a média, utiliza-se
`RSiteSearch("mean")`

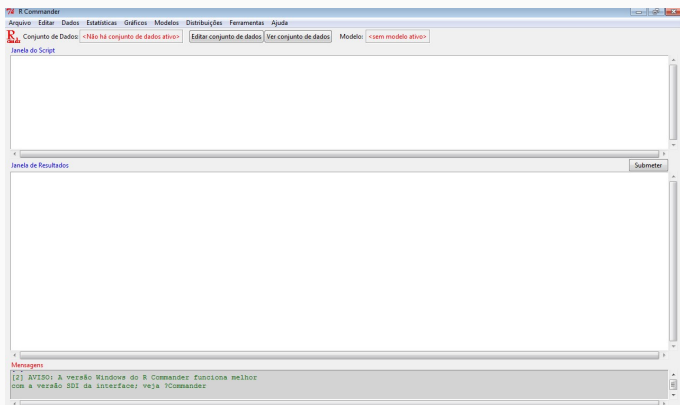
Manuais

- Documentação mais completa está disponível por via electrónica a partir da coleção dos manuais na pagina
<http://CRAN.R-project.org/manuals.html>

R Commander

R COMMANDER

- O R Commander é um pacote complementar para o software estatístico R com interface gráfica.
- Para utilizar o R Commander é necessário carregar o pacote Rcmdr



R COMMANDER

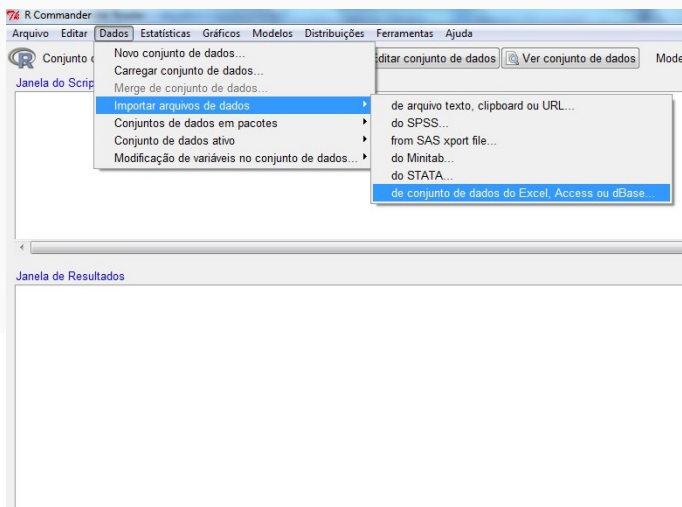
- Existem vários menus na parte superior da janela do R Commander:
 - **Arquivo** - carregar e salvar arquivos de script, para salvar a saída do espaço de trabalho e R, e para sair;
 - **Editar** - editar o conteúdo do script e janelas de saída;
 - **Dados** - contém submenus com os itens do menu para leitura e manipulação de dados;
 - **Estatísticas** - contém submenus com itens para uma variedade de análises estatísticas;
 - **Gráficos** - criar simples gráficos estatísticos;
 - **Modelos** - resumos numéricos, intervalos de confiança, testes de hipóteses, diagnósticos e gráficos para um modelo estatístico;
 - **Distribuições** - contém submenus relacionados com as distribuições de probabilidade;
 - **Ferramentas** - carregar pacotes R alheios ao pacote Rcmdr, e para a fixação algumas opções;
 - **Ajuda** - obter informações sobre o R Commander.

Carregando um Plug in(s)

- Muitos pacotes estão sendo desenvolvidos para integrar o `R Commander`, estes pacotes são chamados de `Plug in(s)`.
- Para carregar um plug in utiliza-se o menu `Ferramentas`
 - `Ferramentas` ⇒ `Carregar plug in(s)` do `Rcmdr`
 - Em seguida escolhe-se o plug in.

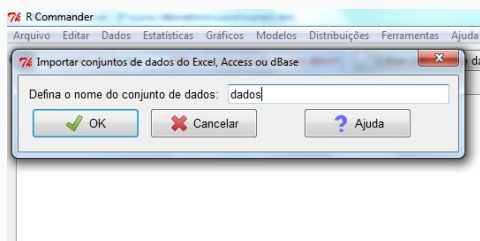
Importar Arquivo de Dados

- Para importar um arquivo de dados utiliza-se o menu:
Dados
 - Dados $\Rightarrow$ importar arquivos de dados
 - Posteriormente escolhe o programa de onde os dados serão importados



Importar Arquivo de Dados

- Em seguida escolhe um nome para o conjunto de dados a ser importado:



- Depois escolhe o diretório e o arquivo que deseja importar.

Manipulando Dados

- Várias manipulações de dados podem ser feitas por meio do Menu `Dados`

- `Dados` ⇒

Modificação de variáveis no conjunto de dados

