S-Plus workshop

7-9 and 14-16 January

students.washington.edu/arnima/s

Statistical software

User-friendly

Excel, SPSS, Statistica

Limited statistical and   Can't automate tasks
graphical functionality

Fast and programmable

Gauss, Maple, Mathematica, Matlab, S, SAS   Visual Basic

Small user group   Limited statistical functionality   Not the   Expensive   Similar speed but
                                                         fastest   but not better   limited statistical and
                                                                                       graphical functionality

Function optimization   **Sheer speed**
ADMB   C++, Fortran

Limited statistical and graphical functionality

Arni Magnusson

7 January 2003

What S offers

Large collection of tools for statistical analysis, constantly updated by a large user community, including leading authorities in statistical fields

Graphics for exploratory analysis and publications

Language for expressing statistical models, object oriented and extensible by users

Arni Magnusson

7 January 2003

The S family

S   Programming language, first version 1976, now 4.0
     Maintained by John Chambers et al., AT&T Bell Laboratories

S-Plus Statistical software that uses S, first version in 1988, now 6.1
     Maintained by Insightful, academic price $115 (1 yr trial is free)
     Choice between GUI and command line interface

R   Statistical software that uses S, first version in 2000, now 1.6.1
     Maintained by R Development Team, free download for anyone
     Command line interface

Arni Magnusson

7 January 2003

Syllabus

Tue 7   **Introduction**
Import data, summarize, regression, plots, export graphs

Wed 8   **Basic statistics**
Descriptive statistics, significance tests, linear models

Thu 9   **Linear models**
Anova, LM, GLM, loess

Tue 14   **Graphics**
Types, multipanel, export graphs

Wed 15   **Data manipulation**
Data objects, describe, extract, sort, manipulate

Thu 16   **Programming**
Functions, import/export, project management, packages

Arni Magnusson
7 January 2003

Today: Introduction

1   **Import data**

2   **Summarize data**

3   **Fit regression model**

4   **Graph model fit to data**

5   **Export graph to Word**

Arni Magnusson

7 January 2003

GUI - Import data

Open Excel
File - Open - mammals.xls
File - Save as - Save as type [CSV]
Close Excel

Open S-Plus
File - Import data - From file - Browse [mammals.csv]
Close data editor
Unselect mammals by clicking on the white space

Arni Magnusson
7 January 2003

GUI - Summarize

Statistics - Data summaries - Summary statistics
Data set [mammals] - Variables [body and brain]
Close report window

Graph - 2D plot - Linear scatter plot
Data set [mammals] - x columns [body] - y columns [brain]

Graph - 2D plot - Log log scatter plot - Graph sheet [GS1]
Data set [mammals] - x columns [body] - y columns [brain]

Arni Magnusson
7 January 2003

GUI - Fit model

Statistics - Regression - Linear

Data set [mammals]

Create formula - Transformation

Select both body and brain - Log - Add

Select log(brain) - Response

Select log(body) - Main effect

Plots - Residuals vs. fit - Untick the "include smooth" option

Switch to the report window

Close the report window

Arni Magnusson

7 January 2003

GUI - Show fitted line

Graph - 2D plot - Log log fit power - Graph sheet [GS1]
Data set [mammals] - x columns [body] - y columns [brain]

Arni Magnusson

7 January 2003

GUI - Export graph

Switch to the graph window, Edit - Copy
Open Word and paste special as picture, to keep the Word file small
Close Word

Arni Magnusson
7 January 2003

First encounter with objects in S

```
                          ls()

                          mammals

                          mammals.lm

                          x

1+8                       ls()

x <- 1+8                  rm(x)

x                         lm

sqrt(x)                   sqrt

                          sqrt(2)

                          lm()

                          ?sqrt

                          ?lm

                          args(lm)
```

Arni Magnusson

7 January 2003

Import data

```
mammals <- read.table("c:/projects/day1/mammals.csv",
                      header=T, sep=",", row.names=1)
```

Arni Magnusson

7 January 2003

Summarize

```
summary(mammals)

plot(mammals$body, mammals$brain)

plot(log(mammals$body), log(mammals$brain))
```

Arni Magnusson

7 January 2003

Fit model

```
mammals.lm <- lm(log(brain)~log(body), data=mammals)

summary(mammals.lm)
```

Show fitted line

```
abline(mammals.lm)
```

Export graph

Switch to one of the graphs, Edit - Copy
(R: File - Copy - Metafile)
Open Word and paste special as picture
Close Word

Arni Magnusson
7 January 2003