ISU.Stat500: An R Package for the Course Stat500 "Statistical Methods"

Yihui Xie [1] Hui Lin [2] Tengfei Yin [3]

November 20, 2009

[1]Department of Statistics, Iowa State University. Email: xie@iastate.edu. This document is based on the course STAT500 taught by Dr Dixon in Fall 2009. We warmly welcome comments on this document.

[2]Department of Statistics, Iowa State University. Email: hlin@iastate.edu.

[3]BCB (Bioinformatics and Computational Biology) Graduate Program, Iowa State University. Email: tengfei@iastate.edu.

Contents

List of Figures

Chapter 1

Comparative Studies

1.1 t-test

1.2 Analysis of Variance

Chapter 2

Linear Regression

2.1 Model diagnostics

2.1.1 Linearity

LOWESS/LOESS is almost equivalent to the linear model[1] when we set the degree of smoothing to be infinitely large. A tutorial we find useful to further understand LOESS is one of the lecture notes by William Jacoby[2], and there is also an intuitive demo in the **TeachingDemos** package (Snow, 2009) which shows the weights actually assigned to each data point for local fitting.

[1] "linear" here is in terms of X variables instead of parameters.
[2] http://polisci.msu.edu/jacoby/icpsr/regress3/lectures/week4/15.Loess.pdf

Chapter 3

Multifactor Studies

3.1 Factorial Designs

3.2 Random Effects

Chapter 4

Misc Functions

We have also collected some functions which are of no direct interest to statistical methods, but they might be helpful for us to save time.

4.1 Dynamic access to course materials

The function *listLinks()* can list all the hyper-links under a web directory with the help of the **XML** package (Temple Lang, 2009). For example, we obtain the names of SAS programs and output files:

```
R> listLinks("http://pdixon.public.iastate.edu/stat500/sas/", "\\.lst$|\\.sas$")
R> ## will get:
R> #  [1] 'http://pdixon.public.iastate.edu/stat500/sas/ancova.sas'
R> #  [2] 'http://pdixon.public.iastate.edu/stat500/sas/bacillus.lst'
R> #  [3] 'http://pdixon.public.iastate.edu/stat500/sas/bacillus.sas'
R> #  ....
R> ## data files are under 'http://www.public.iastate.edu/~pdixon/stat500/data/'
```

We can further download these files automatically using R. There are several examples in the help page; see `?listLinks`.

4.2 Reshaping data between wide and long formats

The function *reshape()* in the **stats** package can reshape data between "wide" and "long" formats. For the rat weight data, one single line of R code is enough for the transformation[1].

```
R> rat = read.table("http://pdixon.public.iastate.edu/stat500/data/ratweight.txt",
+     col.names = c("amount", "type", rep("gain", 10)))
R> rat2 = reshape(rat, varying = list(3:12), idvar = 1:2, direction = "long")
```

[1] Explanations can be found at http://cos.name/en/topic/reshape-the-ratweight-data-into-long-format-stat500 if needed.

Bibliography

Snow G (2009). *TeachingDemos: Demonstrations for teaching and learning*. R package version 2.4, URL http://CRAN.R-project.org/package=TeachingDemos.

Temple Lang D (2009). *XML: Tools for parsing and generating XML within R and S-Plus*. R package version 2.6-0, URL http://CRAN.R-project.org/package=XML.