

# Introduction to R: Homework

Aron C. Eklund      H. Bjørn Nielsen

October 6, 2008

## 1 Exercise 1

### 1.1

There is a really exciting data set included with the default R installation, but I can't remember what it's called. I think it has something to do with *chicken*. Find it by keyword search and load it into your workspace.

### 1.2

Add a new column to `chickwts` that gives the weight in kilograms, rather than grams. Make sure the column has a descriptive name.

### 1.3

Make a box-and-whiskers plot showing the distribution of chicken weight according to feed type. Make sure to label the axes appropriately.

## 2 Exercise 2

### 2.1

Create a numeric vector `x` of length 50 that ranges from  $-\pi$  to  $\pi$ . Create a numeric vector `y1` that is the sine of `x` (in radians). Create a vector `y2` that is the cosine of `x`.

### 2.2

Plot `y1` vs. `x` as a series of points joined by lines. On the same graph, add red-colored points for `y2` vs. `x`. Add a legend.

### 2.3

Plot a histogram of `y2`, making sure there are enough bins to clearly see the trend. Plot a density curve of `y2` using the default parameters. Which plot is more faithful to the true distribution?

## 3 Exercise 3

### 3.1

Point your web browser at <http://data.genome.duke.edu/platinum.php>. Download the "OVC Clinical Data" spreadsheet to your local machine. Now, import the spreadsheet as a data frame into your R workspace, naming the resulting object "clin". Briefly inspect the data.

## 3.2

Confirm that the last three columns are useless, and remove them. Convert the first column to `character` type. Change the name of the third column to "event". Convert the "CA125.POST" and "GRADE" columns into `numeric` values, with the ambiguous entries coded as `NA`s. Fix the "Debulk" column such that only two levels are used ("O" and "S"). Rename the eighth column to "response" and convert it to a `logical` vector (0 = FALSE, 1 = TRUE).

## 3.3

Now that you have "cleaned up" the "clin" object, save it for later use, both as an R object ("clin.rda") and also as a CSV file ("clin.csv").