

228.371 Computer Lab: Multiple Regression

Semester One 2015 - Week 4

Instructions: Read each section and try the commands. Then try the Stream worksheet questions suggested to test your knowledge. The worksheet is “adaptive” which means if you get an answer wrong, you can try again. This quiz is to help you monitor your progress, it does not count toward your mark.

Note that because of fonts, especially for symbols like quotation marks, cutting and pasting commands from this document occasionally will not work - you may have to retype.

1 Multiple Regression

Read the description of the built in R data set `swiss` by typing

```
?swiss
```

Remember you will have to type `data(swiss)` to access this data set. Use the command `pairs(swiss)` to look at scatter plots of each variable pair.

Use the following commands to produce the regression table for predicting `Fertility` with all other variables:

```
m1 <- lm (Fertility~Agriculture+Examination+Education+Catholic+Infant.Mortality, data=swiss)
summary (m1)
```

Are there variables that you would consider removing?

Use the following command to use stepwise variable selection:

```
m0 <- lm (Fertility ~ 1, data=swiss)
m2 <- step(m0, scope=~Agriculture+Examination+Education+Catholic+Infant.Mortality,
          direction="both")
```

Use `plot(m2)` to check the regression assumptions. The final plot will show the Cook's distance contours for 0.5 and 1, if there are any points approaching these values. Typically we are concerned about a point being overly influential if the Cook's distance is larger than 1. You can also compute the Cook's distance for each point with:

```
cooks.distance (m2)
```

Our pairs plot showed that many of our predictors are correlated, so we would like to check the variance inflation factors (VIF). Variance inflation factors show how the inclusion of each variable affects the variances of the other parameter estimates. Typically we should try to remove predictors if the VIF is larger than 10. To compute these factors, we need the `car` library:

```
library (car)
vif (m2)
```

Now try questions 1-6 on the lab worksheet