

LAB 7, MATH 463, SPRING 2014

```
> bfile = url("http://www.uoregon.edu/~dlevin/DATA/bfcars.R")
> load(bfile)
```

Problem 1 (Model Selection). Fit regression models to predict fuel consumption for the data set `bfcars` using

- (1) forward selection
- (2) backwards selection
- (3) stepwise selection
- (4) Find the value of p for which C_p is closest to p .

`T` is a qualitative variable taking the value 1 specifying a manual rather than an automatic gearbox. `G` denotes the number of gears, `C` denotes the number of carburetors. `rar` is the rear-axle ratio, `mt` is the time take to complete a quarter of a mile circuit. `cycls` gives the number of cylinders and `disp` is the car's displacement. To load `bfcars`, use

```
load(url("http://www.uoregon.edu/~dlevin/DATA/bfcars.R"))
```

Problem 2 (Residuals). This problem looks at the population data in `popdata`.

To load `popdata`, use

```
load(url("http://www.uoregon.edu/~dlevin/DATA/population.R"))
```

- (1) Fit a simple linear regression of population against year, and examine the residuals. The residuals should suggest a departure from assumptions.
- (2) Fit a multiple regression model after considering how to fix the failure of the previous model.
- (3) Examine the residuals again. Are there any particular data points that are unusual?
- (4) Tag unusual data points with an indicator variable. Fit a regression model and examine the residuals again.

Problem 3 (Residuals). Consider the data for galapagos island species, `gala`.

- (1) Regress `Species` on all of the covariates.
- (2) Examine the residuals. Do the assumptions of the model seem correct.
- (3) Considering transforming `Species` to fit a model better conforming to the assumptions.
- (4) Examine the residuals and other measures of goodness of fit for the transformed model.