1. The data for this problem are available on the class webpage. The response is a measure of lung function (FEV, in liters). The other variables are the gender of a subject (male/female), the height (in inches),and whether or not that person is a smoker (yes/no). Analyze the data using linear regression (note: there are no interactions in these data, so you don’t need to worry about them), and report your findings. In particular:

(a) Explore the data by plotting them in a meaningful way, and comment.

(b) Fit your statistical models, and select the parameters for your final model.

(c) Make a qq-plot of the residuals. Do they look normal?

(d) Report your findings, and give an interpretation of the parameter estimates for the variables that you used in your final model.

(e) How much of the variability in FEV can you explain with your predictors?

2. For two species of heme, we record the optical density (od) as a function of the H2O2 concentration (conc). The data for this problem are available on the class webpage.

(a) Plot the data.

(b) For each species of heme, fit the model E[od|conc] = b × exp(–c × conc).

(c) Add the curves to the plot.

(d) Test if the lines are identical.

Data Sets:

### 1 ###

dat <- read.csv("http://biostat.jhsph.edu/~iruczins/teaching/140.615/data/fev.csv",row.names=1)

### 2 ###

dat <- read.csv("http://biostat.jhsph.edu/~iruczins/teaching/140.615/data/heme.2species.csv")