STAT 420: Homework #8

Spring 2015, Dalpiaz

Due: April 3, 3:00PM

You may use `R` for all problems on this homework. Please circle or highlight relevant output where possible.

Problem 1

Chemists often use ion-sensitive electrodes (ISEs) to measure the ion concentration of aqueous solutions. These devices measure the migration of the charge of these ions and give a reading in millivolts (mV). A standard curve is produced by measuring known concentrations (in ppm) and fitting a line to the millivolt data. The following data give the concentrations in ppm and the voltage in mV for calcium ISE.

```
ppm <- c(0,0,0,50,50,50,75,75,75,100,100,100,150,150,150,200,200,200)
mV <- c(1.72,1.68,1.74,2.04,2.11,2.17,2.40,2.32,2.33,2.91,3.00,2.89,
        4.47,4.51,4.43,6.67,6.66,6.57)
ises <- data.frame(ppm,mV)
```

(a) Plot the points `mV` (y) versus `ppm` (x).

(b) Fit a simple linear model with `mV` as the response and add it to the plot in part (a). Does a simple linear model seem appropriate here?

(c) Fit a quadratic model with `mV` as the response. Does this model provide a better fit? Explain. (Use a level of $\alpha = 0.05$.)

(d) Add the best fit parabola to the plot from part (a).

Exercise 2

A survey was conducted to study teenage gambling in Britain. (Ide-Smith & Lea, 1988, Journal of Gambling Behavior, 4, 110–118) The data is stored in the data frame `teengamb` (library `faraway`). This dataframe contains the following columns:

`sex`	0 = male, 1 = female
`status`	Socioeconomic status score based on parents occupation
`income`	in pounds per week
`verbal`	verbal score in words out of 12 correctly defined
`gamble`	expenditure on gambling in pounds per year

We will try to model `gamble` as the response and the other variables as predictors.

```
library(faraway)
data(teengamb)
```

(a) Plot `gamble` vs `status`, `gamble` vs `income`, and `gamble` vs `verbal`, using different symbols for males and females. Do these plots suggest the possible need for the interaction terms between sex and the other predictors?

(b) Fit a model with `gamble` as the response and the other variables as predictors that includes the interaction terms between `sex` and the other predictors. What proportion of the variation of `gamble` is explained by this model?

(c) Fit a model with `gamble` as the response and `income`, `sex` and their interaction as the predictors. Use an F-test to determine if you prefer this model, or the model fit in part (b). Use a level of $\alpha = 0.05$.

Exercise 3

Can a corporation's annual profit be predicted from information about the company's chief executive officer (CEO)? Forbes (May, 1999) presented data on company profit (y), (in \$ millions), CEO's annual income (x_1) (in \$ thousands), and percentage of the company's stock owned by the CEO (x_2).

Company	Profit (y, \$)	CEO	Income (x_1, \$)	Stock (x_2, %)
Gap	824.5	Drexler	3,743	1.71
Intel	6,068.0	Grove	52,598	0.13
Gateway 2000	346.4	Waitt	855	43.93
HJ Heinz	746.9	O'Reilly	2,916	1.63
Conseco	630.7	Hilbert	124,579	3.64
Citicorp	5,807.0	Reed	6,200	0.22
Cisco Systems	1,362.3	Chambers	560	0.06
General Electric	9,296.0	Welch	40,626	0.03
America Online	254.0	Case	26,917	0.54
Computer Associates	570.0	Wang	10,614	3.79
Lockheed Martin	1,001.0	Augustine	2,533	0.01
Bear Stearns	538.6	Cayne	23,215	3.44

Source: "Compensation Fit for a King," Forbes, May 1999.

The data can be found in `corporations.csv`.

(a) Fit the interaction model

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + e_i$$

Give the least squares prediction equation and determine whether the overall model is statistically useful for predicting company profit at $\alpha = 0.10$.

(b) Is there evidence to indicate that CEO income x_1 and stock percentage x_2 interact? Use a level of $\alpha = 0.05$.

(c) Based on the least squares estimates of the β parameters, give the estimate of the change in profit for every one thousand dollar increase in a CEO's income when a CEO owns 2% of the company's stock.