

Due in class or in my mailbox before class on April 21

**Section 1 (graded) from Devore, 7th edition:**

The R command  $x \leftarrow -c(1, 2, 4)$  and  $y \leftarrow -c(5, 6, 7)$  followed by  $plot(x, y)$  produces a plot of the points  $(1, 5)$ ,  $(2, 6)$ ,  $(4, 7)$ . This may be helpful in the below. Exercises

12.1.3, 12.1.8, 12.1.10

**Ungraded** Assume an underlying linear regression model with  $\beta_0 = 1$  and  $\beta_2 = 2$ . The idea is to generate “random”  $\epsilon_i$  using R to create “fake data”, and then evaluate an estimate of  $\beta_0$  and  $\beta_1$ . The R commands:

```
x<-rep(c(1:10))
ep<-rnorm(10,0,1)
y<-2*x+1+ep
lm(y~x)
plot(x,y)
```

outputs what you can take to be a blackbox method of estimating  $\beta_0$  and  $\beta_1$  from observations and plots the data  $x, y$ . The intercept coefficient is an estimate of  $\beta_0$  and the “x” coefficient is an estimate of  $\beta_1$ .

a) Repeat the above several times, each time reporting the estimates of  $\beta_0$  and  $\beta_1$ . b) Explain what the R commands are doing (look them up in help files). c) Why do your estimates change? d) The procedure above is sometimes called producing fake data. Why do you think it might have this name?