Math 121 Chapter 13 Solutions Spring 2012

1. (25 pts) Enter the SAT-Math values (x) into list $\texttt{L}_1$ and enter the tuition values (y) into list $\texttt{L}_2$.

   (a) (5 pts) Use $\texttt{LinReg(a+bx) L}_1\texttt{,L}_2\texttt{,Y}_1$. The TI-83 reports that $a = -12.47$ and $b = 0.0556$. So the equation of the regression line is

   $$\hat{y} = -12.47 + 0.0556x.$$

   (b) (3 pts) In the previous part, the TI-83 also reported that $r = 0.2698$, which is the correlation coefficient.

   (c) (3 pts) The value of r is positive and relatively small. That tells us that there is little to no correlation, but what correlation there is, is positive.

   (d) (4 pts) Multiply the slope by the change in x to get the change in $\hat{y}$: $0.0556 \times 10 = 0.556$. We would expect the tuition to increase by \$556.

   (e) (4 pts) Substitute 573 into the regression equation for x and calculate the value of $\hat{y}$. We get $\hat{y} = 19.388$, or \$19,388. You could also enter and evaluate the expression $\texttt{Y}_1\texttt{(573)}$ and get the same answer.

   (f) (3 pts) SAT-Math is the explanatory variable and Tuition is the response variable. (We always let x be the explanatory variable and y be the response variable.)

   (g) (3 pts) These are *confounding* factors.

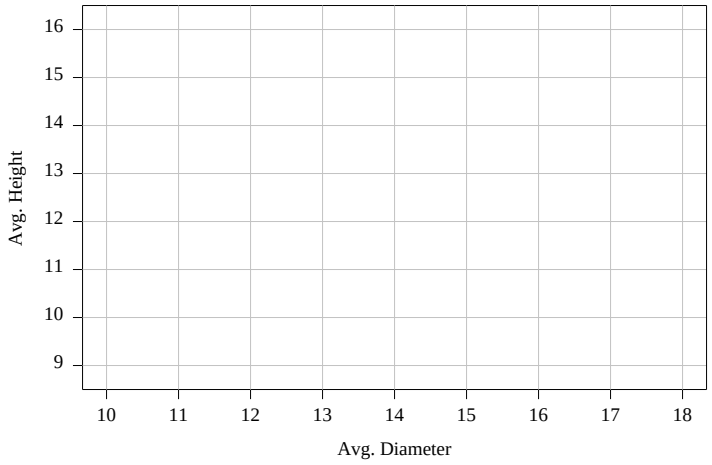
2. (32 pts) The Project on Student Debt recently published a report on student debt after graduation from college in 2007. The report included figures on the average amount of debt by state (average for Virginia is \$18,084) and percentage of students who graduate with debt (percentage for Virginia is 59%). A random sample of 10 states provides the following data.[1]

State	Average Debt (\$1000s) (x)	Percentage with Debt (y)
Alaska	25.0	53
Alabama	20.9	61
Delaware	17.4	50
Indiana	21.3	60
Kansas	18.5	61
Massachusetts	21.1	63
Ohio	22.0	67
Pennsylvania	23.6	71
Rhode Island	23.2	67
South Carolina	20.2	59

[1] `www.projectonstudentdebt.org`

Enter the average debt values (x) into list $\texttt{L}_1$ and enter the percentage values (y) into list $\texttt{L}_2$.

(a) (4 pts) The scatter plot:



(b) (3 pts) Overall, the relationship appears to be positive and strong, with the exception of one value $(25.0, 53)$.

(c) (5 pts) Use $\texttt{LinReg(a+bx) L}_1\texttt{,L}_2\texttt{,Y}_1$. The TI-83 reports that $a = 37.13$ and $b = 1.1289$. So the equation of the regression line is

$$\hat{y} = 37.13 + 1.1289x.$$

(d) (4 pts) The correlation coefficient is $r = 0.4070$, which indicates a moderate to weak positive relationship. (That outlier took its toll on r.)

(e) (3 pts) The coefficient of determination is $r^2 = 0.1657$, which means that 16.57% of the variation in the percentage of students in debt (from state to state) is explained by the variation in average debt (from state to state).

(f) (4 pts) Given that SST = 629.6, find SSR and SSE. Use the formula that SSR $= r^2 \cdot$ SST. We get SSR $= 0.1657(629.6) = 104.3$. Then use the formula that SSE = SST − SSR to get SSE $= 629.6 - 104.3 = 525.3$. (In this problem, I gave the wrong value for SST. It should have been 365.6, in which case SSR = 60.6 and SSE = 305.0.) You could calculate SSR and SSE directly: Compute $\texttt{Y}_1\texttt{(L}_1\texttt{)}$ to get the $\hat{y}$ values and store them in $\texttt{L}_3$. Then compute

$$\text{SSR} = \sum(\hat{y} - \overline{y})^2 = 60.6$$
$$\text{SSE} = \sum(y - \hat{y})^2 = 305.0.$$

(g) (4 pts) Substitute 20.921 into the regression equation for x and get $\hat{y} = 60.75$.

(h) (5 pts) Use the TI-83 function $\texttt{LinRegTTest}$. The seven steps are

1. $H_0 : \beta = 0, \rho = 0.$
$H_1 : \beta \neq 0, \rho \neq 0.$

2. $\alpha = 0.05.$

3. $t = \dfrac{b - 0}{\mathrm{SE}(b)}$, where $\mathrm{SE}(b) = \dfrac{s}{\mathrm{SSX}}$ and $s = \sqrt{\dfrac{\mathrm{SSE}}{n-2}}$.

4. (I'll use the *corrected* values from part (f) so that my answer will agree with `LinRegTTest`.) Compute

$$\begin{aligned} s &= \sqrt{\frac{305.0}{10-2}} = 6.1745 \\ \mathrm{SSX} &= \sum (x - \overline{x})^2 = 47.536 \\ \mathrm{SE}(b) &= \frac{6.1745}{\sqrt{47.536}} = 0.8956 \\ t &= \frac{1.1289 - 0}{0.8956} = 1.261. \end{aligned}$$

5. p-value $= 2 \times$ `tcdf(1.261,E99,8)` $= 0.2430.$

6. Accept H_0.

7. The model is not significant. (It's that outlier again!)

Of course, when you use the `LinRegTTest` function, it will do the calculations in steps 4 and 5 for you. It gives the values $t = 1.260$ and p-value $= 0.2430$.