

- 2.14.a. The regression of FGPA on a constant and SATM can be performed, for instance, by the command `ls fgpa c satm`. The results are in Figure S 2.3. The scatter plot with the regression line indicates that the relation between the two variables is rather weak. The value of s^2 is given by $s^2 = e'e/(n - k) = 2.314/8 = 0.29$.

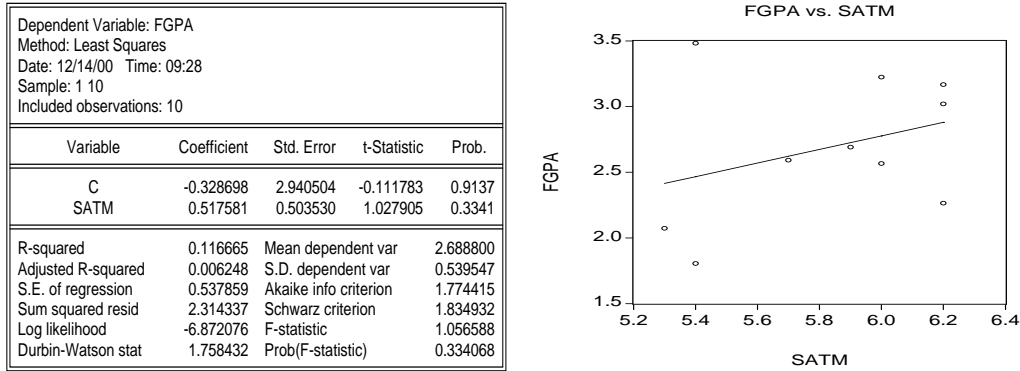


Figure S 2.3: Parts (a) and (b) : Results of regression of FGPA on a constant and SATM

- b. Figure S 2.3 contains the t -statistics, which are -0.11 for a and 1.03 for b . Both indicate that the coefficients are not significantly different from zero at the 5% significance level. The reported P-values are large, both coefficients are not significant
- c. The interval estimate for α has the form $[a - cs_a, a + cs_a]$ with a the estimated value, s_a the standard deviation, and c defined by $P(|t| \leq c) = 1 - \alpha$. Using EViews, c can be calculated as `scalar ct=@qtdist(0.975,8)`. We find $c = 2.306$ and the interval estimates

$$\text{for } a : [-0.329 - 2.306 \cdot 2.941, -0.329 + 2.306 \cdot 2.941] = [-7.111, 6.453],$$

$$\text{for } b : [0.518 - 2.306 \cdot 0.504, 0.518 + 2.306 \cdot 0.504] = [-0.644, 1.680].$$

- d. The estimated regression equation is

$$\text{FGPA} = -0.329 + 0.518 \cdot \text{SATM} + e,$$

where e denotes the residuals of this model. For $\text{SATM}=6$, the best prediction is $\text{FGPA} = -0.329 + 0.518 \cdot 6 = 2.779$. For the estimated variance of the forecast we use the formula of Section 2.4.1,

$$s_f^2 = s^2 \left(1 + \frac{1}{n} + \frac{(x_{n+1} - \bar{x})^2}{\sum (x_j - \bar{x})^2} \right).$$

From the statistics in Table S 2.7 we obtain

$$n = 10$$

$$\bar{x} = 5.83$$

$$s_x = \sqrt{\sum (x_i - \bar{x})^2 / (n - 1)} = 0.356, \text{ so that } \sum (x_i - \bar{x})^2 = 1.141.$$

The standard error of the regression is reported in Figure S 2.3 as 0.538, so that

$$s_f^2 = (0.538)^2 \left(1 + \frac{1}{10} + \frac{(6 - 5.83)^2}{1.141} \right) = 0.326.$$

(Using the forecast function of EViews, the standard deviation s_f of the forecast is given as 0.570569, leading to $s_f^2 = 0.326$). The prediction interval is given by

$$[\hat{y} - cs_f, \hat{y} + cs_f] = [2.779 - 2.306 \cdot 0.571, 2.779 + 2.306 \cdot 0.571] = [1.462, 4.096].$$

Date: 12/14/00 Time: 10:58		
Sample: 1 10		
	SATM	FGPA
Mean	5.830000	2.688800
Median	5.950000	2.642000
Maximum	6.200000	3.482000
Minimum	5.300000	1.805000
Std. Dev.	0.356059	0.539547
Skewness	-0.383898	-0.151034
Kurtosis	1.578375	1.936287
Jarque-Bera	1.087720	0.509471
Probability	0.580503	0.775122
Observations	10	10

Table S 2.7: Part (d) : Statistics on SATM and FGPA

- e. The whole prediction is built on the results of the regression. Therefore all of the Assumptions 1-7 should be met, strictly (to have a strict interpretation) or at least approximately (so that the results hold approximately true). For our data, the estimates of a and b are not significant and the prediction interval is very wide, as it runs from 1.5 to 4.1. The value of SATM=6 falls well inside the range of values of the 10 students in the table in Exercise 1.11, and the prediction interval in part (d) contains all the 10 observed FGPA scores in this table (these FGPA scores range between 1.8 and 3.5). So, in this case we may be quite confident about the constructed interval.