ISyE 4031 Regression and Forecasting
Homework 2 Solutions
Spring 2016

1. Exercise 3.3.

a. The slope estimate, $b_1 = 5.70657$ may be interpreted as follows: An increase of one point in the GPA corresponds to an increase of $5,706.57 in the average starting salary.
The intercept estimate, $b_0 = 14.8156$ clearly has no practical interpretation because it indicates a mean starting salary of $14,815.60 when the GPA= 0.000.

b. $\hat{y} = 14.8156 + 5.70657(3.25) = 33.362$ ($33,362).

c. Calculations:

$$SS_{xy} = \Sigma x_i y_i - \frac{(\Sigma x_i)(\Sigma y_i)}{7} = 709.372 - \frac{(21.57)(226.8)}{7} = 10.5040$$

$$SS_{xx} = \Sigma x_i^2 - \frac{(\Sigma x_i)^2}{7} = 68.3071 - \frac{(21.57)^2}{7} = 1.8407$$

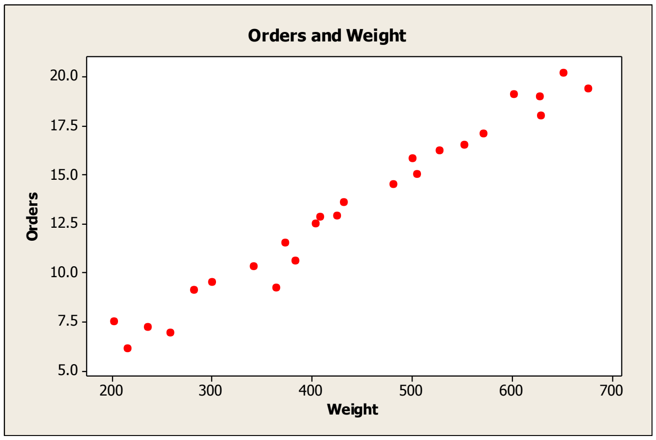
$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{10.5040}{1.8407} = 5.7065$$

$$\bar{x} = \frac{\Sigma x_i}{7} = \frac{21.57}{7} = 3.0814 \qquad \bar{y} = \frac{\Sigma y_i}{7} = \frac{226.8}{7} = 32.4$$

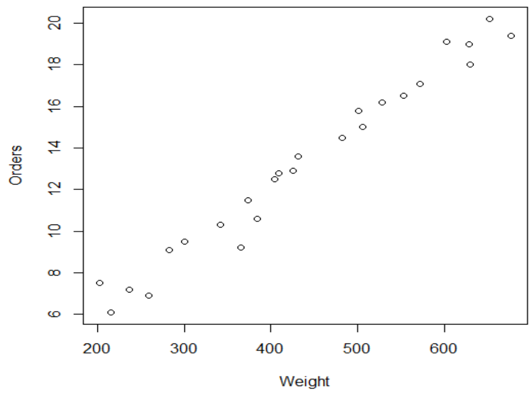
$$\hat{\beta}_0 = \bar{y} - b_1\bar{x} = 32.4 - (5.7065)(3.0814) = 14.816$$

Note: For more accurate estimate carry more decimal places.

2. a. Yes, it seems that a straight-line model is very appropriate.



(Minitab output)



(R output)

b. The solution to the simple linear regression model is Orders = 0.191 + 0.0297Weight.

Minitab output:

```
The regression equation is
Orders = 0.191 + 0.0297 Weight
```

```
Predictor       Coef    SE Coef       T      P
Constant      0.1912     0.4747    0.40  0.691
Weight      0.029703   0.001030   28.82  0.000

S = 0.725784   R-Sq = 97.3%   R-Sq(adj) = 97.2%

Analysis of Variance
Source          DF      SS      MS       F      P
Regression       1  437.64  437.64  830.82  0.000
Residual Error  23   12.12    0.53
Total           24  449.76
```

R output:
Residuals:

Min	1Q	Median	3Q	Max
-1.83270	-0.22077	0.05544	0.46039	1.27914

Coefficients	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	0.19122	0.47475	0.403	0.691
Weight	0.02970	0.00103	28.824	<2e-16 ***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Residual standard error: 0.7258 on 23 degrees of freedom
Multiple R-squared: 0.9731, Adjusted R-squared: 0.9719
F-statistic: 830.8 on 1 and 23 DF, p-value: < 2.2e-16

Analysis of Variance Table
Response: Orders

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Weight	1	437.64	437.64	830.82	< 2.2e-16 ***
Residuals	23	12.12	0.53		

c. From the output:
SSE = 12.12, $s^2 = MSE$ = 0.53, and s = 0.7258.

3. $\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})=\sum_{i=1}^{n}(x_iy_i - x_i\bar{y}-\bar{x}y_i+\bar{x}\bar{y})$

$$= \sum_{i=1}^{n}x_iy_i - \bar{y}\sum_{i=1}^{n}x_i - \bar{x}\sum_{i=1}^{n}y_i + \sum_{i=1}^{n}\bar{x}\bar{y} =. \sum_{i=1}^{n}x_iy_i - \bar{y}\sum_{i=1}^{n}x_i - n\bar{x}\bar{y}+n\bar{x}\bar{y}$$

$$= \sum_{i=1}^{n}x_iy_i - \frac{\sum_{i=1}^{n}y_i\sum_{i=1}^{n}x_i}{n}.$$

4. a. $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

$$SS_{xx} = 143215.8 - \tfrac{1478^2}{20} = 33991.6$$

$$SS_{xy} = 1083.67 - \tfrac{(1478)(12.75)}{20} = 141.445$$

$$SS_{yy} = 8.86 - 20(0.6375^2) = 141.445$$

$$\hat{\beta}_1 = \frac{SS_{xy}}{SS_{xx}} = \frac{141.445}{33991.6} = 0.00416$$

$$\hat{\beta}_0 = \tfrac{12.75}{20} - (0.0041617512)(\tfrac{1478}{20}) = 0.32999$$

$$\hat{y} = 0.32999 + 0.00416x\,.$$

b. $SSE = SS_{yy} - \hat{\beta}_1\ SS_{xy} = 0.731875 - 0.00416(141.445) = 0.1435$

$$\hat{\sigma}^2 = MSE = \frac{SSE}{n-2} = \frac{0.1435}{18} = 0.0079\,.$$

c. $\hat{y} = 0.32999 + 0.00416(85) = 0.6836$.

d. $\hat{y} = 0.32999 + 0.00416(90) = 0.7044$.

e. Residual = Prediction error = $e_i = y_i - \hat{y} = 0.6 - 0.7044 = -0.1044$.

f. $\hat{\beta}_1 = 0.00416$.