Name: *Caleb McWhorter — Solutions*
MAT 222
Spring 2019
Homework 8

"Yeah, Mr. White! Yeah, Science!"
–Jesse Pinkman, Breaking Bad

Problem 1: A research group is trying to predict the average amount of hours it takes to fully 'adapt' to a new work environment using the number of minutes spent in work training, the amount of minutes spent in computer training, and the amount of time spent reviewing orientation materials.

Analysis of Variance

Source	DF	Adj SS	Adj MS	F-Value	P-Value
Regression	*3*	10037.1	3345.7	*19.19*	0.000
Train	*1*	*4102.8*	4102.8	23.53	0.000
Computer	*1*	6259.8	*6259.8*	35.91	0.000
Review	*1*	806.5	806.5	4.63	0.036
Error	*50*	8716.6	*174.3*		
Total	53	18753.7			

Model Summary

S	R-sq	R-sq (adj)	R-sq (pred)
13.2035	*53.52%*	50.73%	0.00%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	25.14	4.93	*5.10*	0.000	
Train	*-0.02571*	0.00530	-4.85	0.000	7.89
Computer	0.03137	0.00523	5.99	*0.000*	8.19
Review	-0.891	0.414	-2.15	0.036	1.15

The regression equation is

Adapt $= 25.14 - 0.02571$ Train $+ 0.03137$ Computer $- 0.891$ Review

(a) Fill in the missing entries above.

(b) What is the average adjustment time for someone that spent 1.5 hours in training, 10 hours in computer training, and spent 30 minutes reviewing orientation materials?

$$\textit{Adapt} = 25.14 - 0.02571(90) + 0.03137(600) - 0.891(30) = 14.92 \textit{ hours},$$

i.e., 14 hours and 55 minutes.

(c) What is the correlation coefficient for this model?

$$r = \sqrt{r^2} = \sqrt{0.5352} = 0.7316$$

(d) What was the total number of subjects examined to create this model?

We have DFT= 53 *and DFT=* $n-1$ *so that* $n = 54$.

(e) Construct a 95% confidence interval for β_2.

We have $n = 54$ *and 3 variables so that gives degrees of freedom* $n - p - 1 = 54 - 3 - 1 = 50$*, the degrees of freedom of the error (DFE). This gives* $t^* = 2.009$. *We know that* β_2 *is the variable computer and that* $b_2 = 0.03137$ *with standard error* $SE_{b_2} = 0.00523$. *Therefore, we compute* $b_2 \pm t^* SE_{b_2} = 0.03137 \pm 2.009(0.00523)$ *to find confidence interval* $(0.021, 0.042)$

(f) Find the value of $\sum(x_i - \overline{x})^2$ for this data.

We have $SE_{b_2} = \dfrac{s}{\sqrt{\sum(x_i - \overline{x})^2}}$. *But then we have* $\sum(x_i - \overline{x})^2 = \dfrac{s^2}{SE_{b_2}^2}$. *Then*

$$\sum(x_i - \overline{x})^2 = \frac{13.2035^2}{0.00523^2} = 6373452.6.$$

(g) Perform the F-test for this model. State your null and alternative hypotheses, F-statistic, degrees of freedom of the numerator/denominator, p-value, and conclusion.

We have hypotheses

$$\begin{cases} H_0 : \beta_1 = \beta_2 = \beta_3 = 0 \\ H_a : \text{ not all } \beta_i = 0 \end{cases}$$

We already found F*-statistic* 19.19 *with* p*-value* 0.000. *[Note the degrees of the freedom of the numerator is* $p = 3$ *and the degrees of freedom of the denominator is* $n - p - 1 = 50$*.] Therefore, we reject the null hypothesis, i.e. the model is statistically significant.*