

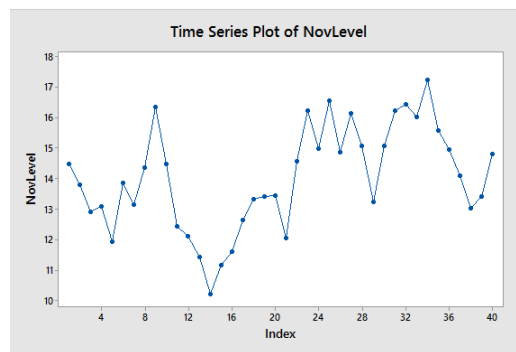
Homework 12

Roly Vicaría
STAT501 Fall 2015

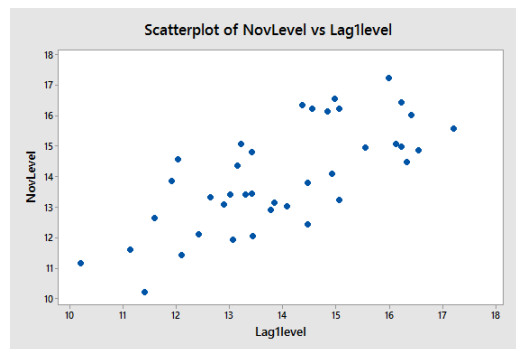
December 9, 2015

0.0.1 Question 1

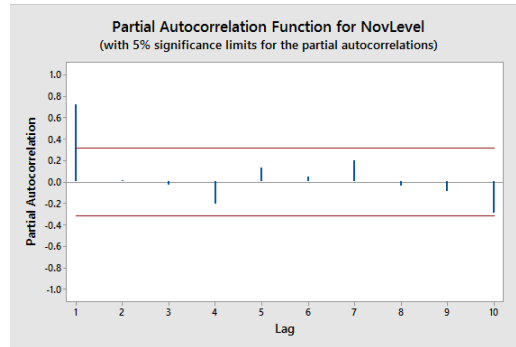
- a) There appears to be a slight upward trend in the data.



- b) The plot of NovLevel vs the Lag1level does show a positive, linear pattern suggesting that the first-order autoregression model could be useful.

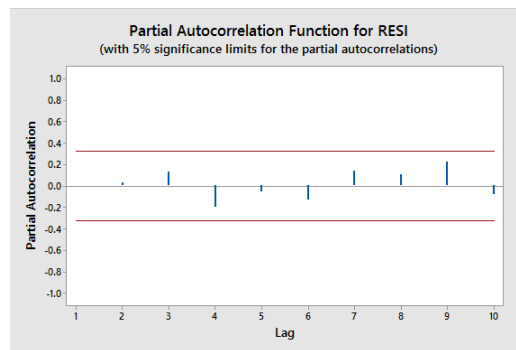


- c) The partial autocorrelation shows a significant spike at a lag of 1 and much lower spikes for the subsequent lags. It indicates that a AR(1) model would likely be feasible for this data set.



- d) i. $\text{NovLevel} = 3.88 + 0.723 \text{ Lag1level}$
 ii. 14.5841

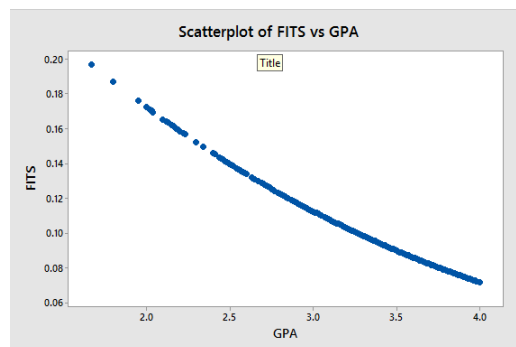
- e) The partial autocorrelation shows only weak correlations, so the residuals could reasonably be assumed to be independent of each other.



0.0.2 Question 2

a)
$$\hat{p} = \frac{e^{-0.573 - 0.498(GPA)}}{1 + e^{-0.573 - 0.498(GPA)}}$$

- b) The plot shows that the probability of cheating on an exam decreases as the GPA increases.



- c) The odds ratio for GPA is 0.6076. This can be interpreted as saying that for every 1 point increase in GPA, the odds of having cheated on an exam are multiplied by 0.6076.

$$d) \hat{p} = \frac{e^{-0.573 - 0.498(3.0)}}{1 + e^{-0.573 - 0.498(3.0)}} = 0.1266$$

$$e) \hat{p} = \frac{e^{-0.573 - 0.498(4.0)}}{1 + e^{-0.573 - 0.498(4.0)}} = 0.07692$$

$$Odds = \frac{p}{1-p} = \frac{0.07692}{1-0.07692} = 0.0833$$

f) In the expanded tables output, we can see the Z-value and p-value for the SkipClass term which shows a p-value of 0.001 indicating a significant predictor.

g) The odds ratio for SkipClass is 1.4138. This can be interpreted as saying that for every additional class skipped by a student in a typical week, the odds of having cheated on an exam are multiplied by 1.4138, assuming GPA is held constant.

$$h) \frac{p}{1-p} = e^{-1.190 - 0.375(3.5) + 0.346(1)} = 0.1157$$

$$i) 0.1157 * 1.4138 = 0.1636$$

j) Of the 4 additional categorical predictors, it looks like SmokeCig and FakeID are significant predictors of ever having cheated on an exam with p-values of .005 and .010, respectively. ChtdSO is on the fence since its p-value is less than 0.1, but I will not consider it significant here. The 4th predictor, SmokeMJ, has a very high p-value of .412 and therefore not a significant predictor of having cheated on an exam.

The odds ratio of SmokeCig is 2.2769 which says that the odds of ever having cheated on an exam for students who smoke cigarettes is 2.2769 times higher than that of students who do not smoke cigarettes (all else being constant).

The odds ratio of FakeID is 1.9454 which says that the odds of ever having cheated on an exam for students who have ever used a fake id to obtain alcohol is 1.9454 times higher than that of students who have never used a fake id (all else being constant).

0.0.3 Question 3

$$a) Falls = e^{0.489 - 1.069(Inter) - 0.047(Gender) + 0.00947(Balance) + 0.00857(Strength)}$$

$$b) H_0 : \beta_{Gender} = 0$$

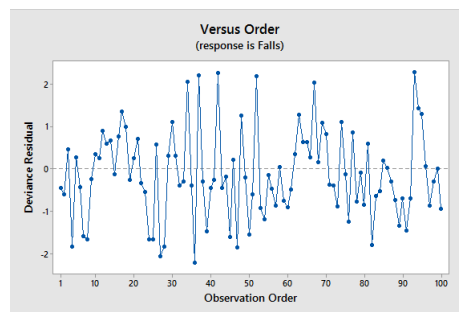
$$H_a : \beta_{Gender} \neq 0$$

$$\text{Test statistic: } G^2 = 108.941 - 108.790 = 0.151$$

$$p\text{-value} = 0.698$$

Based on the p-value, we do not have enough evidence to reject the null hypothesis, therefore we conclude that $\beta_{Gender} = 0$.

c) The Residuals vs order plot does not reveal any outlying cases, values are generally in the -2 to 2 range.



d) $Falls = e^{0.444 - 1.078(Inter) + 0.00947(Balance) + 0.00898(Strength)}$

Based on this estimate equation, aerobic exercise does reduce the frequency of falls when controlling for balance and strength.

e) Based on the Pearson and Deviance statistics' p-values, 0.243 and 0.173, respectively, there is no evidence of lack-of-fit.

$$R^2 = 1 - \frac{-2\ell(\hat{\beta})}{-2\ell(\hat{\beta}_0)} = 1 - \frac{108.941}{199.194} = 0.4531$$