Syed Mahmood

Lab 10 Writeup

Firstly, thank you for being accommodating about missing lab due to being sick.

Part 1 [please see output for all regression statistics]:

The RESET command is only done via autoreg, so what we had to do is take the predicted value, then create two variables to square and cube it, then re-run our regression with these variables. I named these values py2 and py3 respectively. Both of these have very low t-values and high pr values, meaning they are not statistically significant. When we run autoreg, we get reset values to the power of 2, 3, and 4 respectively. We also ran the Normal test which is a high value at 566.3407with low probability for error. Essentially, the reset command in autoreg does this automatically whereas what we did in the second regression was done manually. Parameter estimates for the four variables are the exact same in autoreg as they are in the first proc reg command.

In the RESET test, we can also see the pr value go up as the power increases. This trend matches what we had in our second regression when we did it manually.

Overall we also have a higher r-squared value in the autoreg command regression over the original regression we ran.

Part 2 [please see output for all regression statistics]:

Distance can never be negative, so we expect log(dist) to be a positive value as it is a log of a positive value. We get a parameter estimate of a positive value at .36 with a t-value of 5.55 with a low Pr value, meaning this value is stasticially significant.

Once we re-do the regression with more logged variables, we get a parameter estimate of .055, a lower ldist value at .05 with a lower t-value and a much higher Pr value at .33, meaning this is not statistically significant.

When we add [log(inst)]^2 to the regression, we see the parameter estimate at .1853 with a t-value of 2.97 with a low pr value of .0035, meaning that this test is statistically significant.

In our final regression, however, adding [log(dist)]^2 to the regression, we see that our standard error tremendously increases, our parameter estimate is now .8696, our t-value drops to .42 and our Pr value becomes very high at .67 meaning there is a lot of room for error. This means that adding [log(dist)]^2 to our regresson has resulted in our ldist value becoming not statistically significant at all.