# MAT 303 Module Three Problem Set Report

Second Order Models

Benjamin Leanna

Benjamin.leanna@snhu.edu

Southern New Hampshire University

This week I’ll be taking the role of an analyst working for the government. I have access to a set of historical data that can be used to study wage growth of the labor force. The results I find will be useful for governments interested in wage growth patterns based on their economic agenda. I will be making visualizations with scatterplots to show wage growth to inflation, looking at intervals between predictions and confidences, and making different multiple regression models with an array of distinct variables.

The important set of variables I’ll be using for this data set include wage growth rate, rate of inflation, unemployment rate, GDP growth rate, and whether the economy is in a recession or not in recession. In this data set there are 99 samples (rows) with 6 separate variables (columns).

Chart, scatter chart

Description automatically generated For this first part, I will be creating a scatterplot of wage growth and inflation, as seen here:

As you can see from the scatter plot, the wage growth slows down at a higher rate as inflation grows. This would indicate to me that a second order model would appropriate here for the analysis. The general form of this equation would be  **which would make mine .**

![A screenshot of a computer

Description automatically generated with low confidence]()

**From this table, the is 0.8614 while the adjusted is 0.8585. What this tells me about the is that this model is approximately 86.14% confident in its accuracy of fitting the terms to a curve or line. The adjusted value is always lower than the original since it adjusts for the number of terms in the model and is a bit more precise. The reason we look at both numbers is because if you were to compare two different models with different amounts of variables, you would use the adjusted value. The beta estimates for inflation (1.81077) and I() (-0.08129) will reflect how adding 1 unit of either would affect the dependent variable of wage growth. For every 1 unit of inflation, wage growth will rise by inflation’s beta estimate, and for every unit in I(), wage growth will lower by I()’s beta estimate.**

**When evaluating this model at a 5% level of significance, I want to look at the null hypothesis, the alternative hypothesis, and the P-values within the model. The null hypothesis here would be and the alternative hypothesis being . With a 5% level of significance, the P-value of either inflation or I() should be under 0.05 to be found significant. With both P-values ( and ) being well under 0.05, both variables are significant. The overall F-test’s P-value is making the test of the model itself significant.**

**Using this model to make a wage growth prediction when inflation is 7.41 would use the algorithm and would equate to a positive 8.5593 wage growth.**

![A screenshot of a computer

Description automatically generated with medium confidence]()**The prediction interval for this model has a lower end of 6.4568 and an upper end of 10.6616. With the model itself showing an 86.14% confidence in its accuracy, it comes out with a higher range than the 95% confidence interval. This confidence interval would be 8.2473 on the lower end and an 8.871 on the upper end.**

**For the next part, I want to write a general form of the complete second order regression model using wage growth as the response variable, and inflation and GDP growth as predictor variables. The general form would be . In this model, is wage growth, is inflation, and is GDP growth. From my regression model, the algorithm becomes . The values of and adjusted are 0.9113 and 0.9065, respectfully. This tells me that the model has a 91.13% confidence in accuracy and 90.65% confidence in accuracy when considering the number of variables present. The beta numbers for are -0.003176 and for are -0.027371. The and are added into this model as their own independent variables as so to make up a quadratic model and further the accuracy.**

**Looking at this model to see if it is significant at the 5% level of significance, I will again identify the null hypothesis, the alternative hypothesis, and P-value to make the conclusion. The null hypothesis here being and the alternative hypothesis being The conclusion of the overall F-test has a P-value of making this model significant at the 5% level of significance. Looking at the individual variables however, the only two that fall below 0.05 to be significant are inflation ( and (.**

**Making a prediction using this regression model when inflation is 7.41 and GDP growth is 9.59 would have an equation that looks like this: . With these numbers the prediction comes to an 8.45793 wage growth. The prediction interval for this has a lower end of 6.744 and upper end of 10.1718 while the confidence interval has a lower end of 8.1751 and upper end of 8.7407. Again, the prediction interval has a wider range that the confidence interval since the model itself leaves for more than 5% rate of error approximation while the confidence interval is calculated exactly at a 5% rate of error** ![A screenshot of a computer

Description automatically generated with medium confidence]()**approximation.**

On this next model, I’ll be writing the general form for a complete second order regression model for wage growth using inflation and economy as predictor variables. The general form is  **with being wage growth, being inflation, and being whether or not the economy is in a recession (indicated with a 1 for yes or 0 for no). My model would then be . The and adjusted values are 0.8738 and 0.867, respectfully. This tells me that from the overall F-test, the model is 87.38% confident in predicting the correct wage growth given the independent variables. The adjusted value is a bit lower because it considers the number of variables in the model as well.**

**At a 5% level of significance, the overall F-test of the model would be significant with its P-value being . The null hypothesis here is and the alternative hypothesis being . With the models P-value we can dismiss the null hypothesis and go with the alternative hypothesis instead. In the individual T-tests of this model only one independent variable was below the 0.05 threshold. This variable being with a P-value of 0.00167. Since this variable is significant, it also means that the inflation variable is significant as well, regardless of its P-value. Using this model for a prediction of wage growth If inflation is 7.41 and the economy is not in a recession would like this: This equates to an 8.438 wage growth. The 95% prediction interval has a lower end of 6.3961 and an upper end of 10.4796 while the confidence interval has a lower end of 8.1175 and an upper end of 8.7582. Again, the confidence interval has a narrower range since it is accuracy approximation is 95% while the prediction interval from the model is much lower (87.38%).**

**Concluding this report and based on the analysis that I’ve performed and assuming that the sample size is sufficiently large, I would indeed recommend using this model. The reason being is that it has a high (80%+) confidence in its accuracy and a P-value close to 0 (. The practical importance of the analyses that were performed here are to not only measure correlation between specific variables to see how each one affects the outcome alone, but also to see how they interact with each other to affect a response. With the addition of using the quadratic model, we can get a more precise reading on the rate of change between different variables as well. We can see where a model starts to shift at an inflection point to where a positively added variable starts to become detrimental.**