# MAT 303 Module Three Problem Set Report

Second Order Models

Luisa Heisler

luisa.heisler@snhu.edu

Southern New Hampshire University

## **1. Introduction**

The data set I will be exploring is on the economy and the variables that can influence it. These results could be used to study and predict how different variables can affect the economy. The variables include wage growth, inflation, unemployment rate, and economy (state of). Economists could create models with different scenarios to see how changes in certain variables would affect other variables. Specifically, how the economy is impacted by historical wage growth patterns. The analyses being run in this problem set include three quadratic (second order) multiple regression models. One with a quantitative variable, one with two quantitative variables and one with a quantitative and qualitative (categorical) variable.

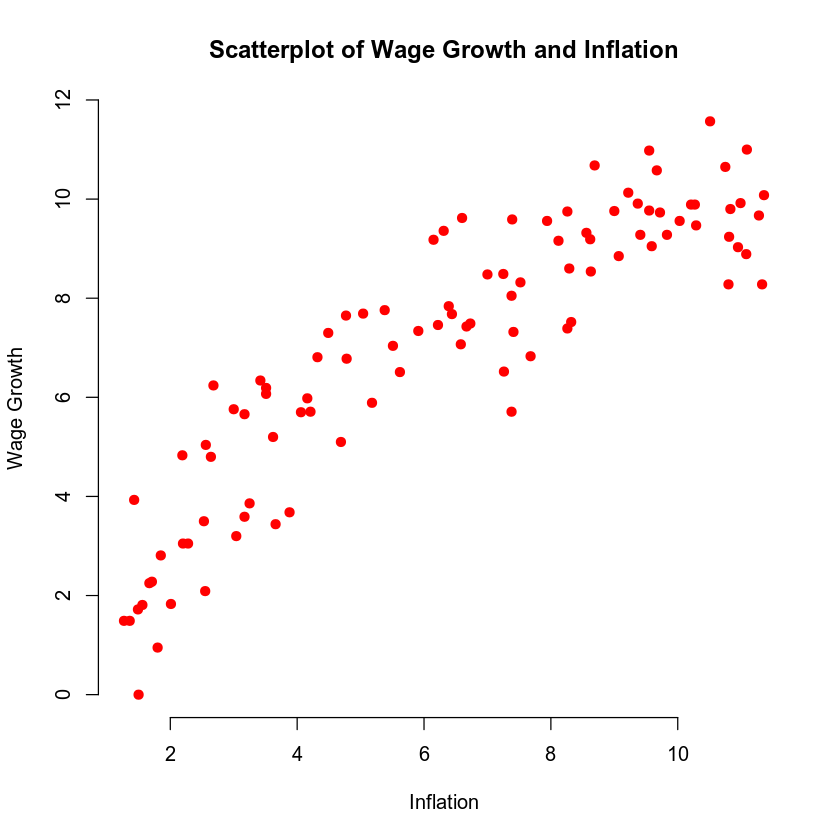
## **2. Data Preparation**

The important variables that I will be analyzing in this problem set are as follows, wage growth as response variable and inflation as predictor variable. Wage growth as response variable, and inflation and GDP growth as predictor variables. Wage growth as response variable, and inflation and economy as predictor variables. There are 6 columns in the full data set, for different variables relating to the economy and 99 rows.

## **3. Quadratic (Second Order) Model with One Quantitative Variable**

### Correlation Analysis

A scatter plot of wage growth and inflation was created (shown below). Looking at the scatterplot, posted below, it shows that the wage growth rate is non-linear related to inflation. As inflation increases, wage growth increases and then tends to flatten.



A second order model is the appropriate move in this case since the predictor variable has a non-linear relationship with the response variable. The predictor variable can have a trend that increases or decreases based on factors and it is not only dependent on the response variable.

### Reporting Results

The general equation for the second order regression model for wage growth using inflation as the predictor variable is:

E (Y) = β0 + β1X + β2X2

β0 is the intercept. β1 is the popular regression term for inflation and β2 is the popular regression term for inflation2(inflation squared). The X and X2 terms are the values of inflation and inflation2 that can have hypothetical terms substituted for it.

The model equation is:

ŷ =0.7133 + 1.1144 x− 0.0326 x2

The value of R2 (R-squared) is 0.8338 and the value of R (adjusted R-squared) is 0.8303. The Ra2 value is the coefficient of multiple determination and shows that approximately 86% of the variance in wage growth can be explained using a model that uses inflation as a predictor.

### Evaluating Model Significance

The null and alternative hypothesis for an overall F-test are:

H: β=β=0012

Ha: at least one β≠0fori=1,2ai2.2e-16

The P-value of the overall F-test is which is significantly lower than the 5% level of significance. As such, the null hypothesis should be rejected in favor of the alternative hypothesis. With this it can concluded that a statistically significant relationship exists between wage growth and inflation.

To see which terms are significant in the model, an individual T-test is used for each variable. The null and alternative hypotheses for these are:

H: β=0for somei=1, 2,...,n

Ha: β1≠0

The P-value for inflation is 2e-16 and the P-value for inflation2 is 8.81e-09. Both are well below the 5% level of significance. With these P-values, it can be concluded that both variables, individually, are statistically significant at a 5% level of significance.

### Making Predictions Using Model

If we want to find the predicted wage growth if inflation is 2.54, we would want to use the equation for second order regression model.

E(Y)= β0 + β1X + β2X2

Y= 0.7132 + 1.1144(2.54) – 0.0326(2.54^2)

Y= 0.7132 + 2.8306 – 0.2103

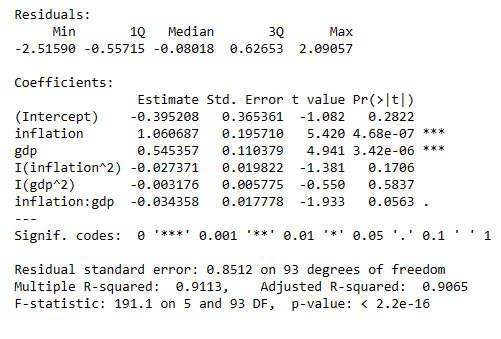
Y= 3.3335

## **4. Complete Second Order Model with Two Quantitative Variables**

### Reporting Results

Our next model solving for wage growth we are going to run includes inflation as the independent variable like before, but we are also going to include GDP growth. The general form of our model is: E (Y) = β0+β1x1+β2x2+β3x2/1+β4x2/2

where y is wage growth, x1 is inflation, x2 and is GDP.



So, my model equation is:

E(Y)= -0.3952 + 1.0606 x1 + 0.5453 x2 – 0.0273 x2/1 – 0.0031 x2/2

Our R2 for this equation is .9113 telling us that 91.13% of the variation in the response variable is explained by the model that uses these predictors and the interaction term. Our adjusted R2 has a value of 90.65%. The adjusted R2 considers the number of predictors in the model it is adjusted downward slightly. This is a very large value for our coefficient of determination letting us know that these variables play a major role in the movement of wage growth.

### Evaluating Model Significance

We are going to be using the same null hypothesis and alternative hypothesis for our t-test that we did on the previous model. The null hypothesis is claiming that not one of the variables has value, or that their slope will be zero and the alternative is stating that at least one of the variables use will have value.

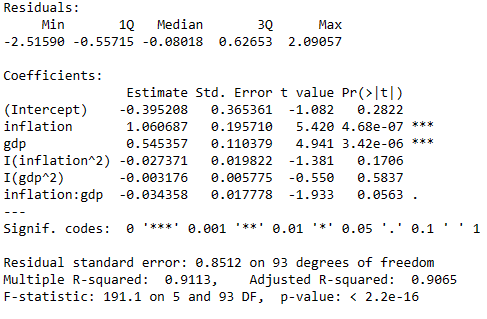
Null Hypothesis: H0: β1 β2 = 0

Alternative Hypothesis: Ha: At least one B1 ≠ 0

Looking at the p-value of our F-statistic we can see that it is close to zero as 2.2 e-16. This is well below our 5% level of significance and tells us we can reject the null hypothesis and accept the alternative hypothesis that at least one of the variables has value. We will run the individual t-tests in order to see which one. The null hypothesis will look at the p-value for each variable and if it is greater than 5% level of significance then we would state that variable has a slope equal to zero or has no significance to the model. The alternative hypothesis looks at each variable separately and claims that it does have value.

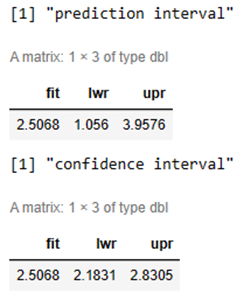
Null Hypothesis: H0: Bi = 0

Alternative Hypothesis: H0: Bi ≠ 0



All of them are well below the 5% level of significance and we would reject the null hypothesis and claim that all of them are valuable to our model.

### Making Predictions Using Model



If we want to find the predicted wage growth if unemployment is 2.50 and GDP growth is 6.50, we would want to use the equation for second order regression model.

Y=−0.39508+1.060687(2.50) +0.545357(6.50) −0.034358(2.50∗6.50) −0.027371(2.50^2)−0.003176(6.50^2)

Y=5.0721

The 95% prediction interval for wage growth is the prediction interval tells us that we are in the bounds of being 95% certain that wage growth will lay within these lower and upper bounds (only if unemployment is 2.50 and GDP growth is 6.50.

The 95% confidence interval for wage growth is the confidence interval tells us that are in the bounds of being 95% certain that wage growth will be within these upper and lower bounds only if unemployment is 2.50 and GDP growth is 6.50.

## **5. Complete Second Order Model with One Quantitative and One Qualitative Variable**

### Reporting Results

*Report the results of the regression model. Address the following questions in your analysis:*

* *Write the general form and the prediction equation of the complete second order regression model for wage growth using unemployment and economy as predictor variables.*
* *Create this second order regression model for wage growth using unemployment and economy as predictors.* *Write the prediction model equation using outputs obtained from your R script.*
* *What are the values of R-Squared (R-squared) and Adjusted R-Squared (Adjusted R-squared) for the model? Provide your interpretation of these statistics.*

Caution sign icon Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include R code in your report.

### Evaluating Model Significance

*Evaluate model significance for the regression model. Address the following questions in your analysis:*

* *Is the model significant at a 5% level of significance? Carry out the overall F-test by identifying the null hypothesis, the alternative hypothesis, the P-value, and the conclusion of the test.*
* *Which terms are significant in the model based on individual T-tests? Use a 5% level of significance.*

Caution sign icon Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include R code in your report.

### Making Predictions Using Model

*Make predictions using the regression model. Address the following questions in your analysis:*

* *What is the predicted wage growth if unemployment is 2.50 and the economy is* ***not*** *in recession? (be sure to use single quotes when setting the value for economy)*
* *What is the 95% prediction interval for the wage growth? Interpret the interval.*
* *What is the 95% confidence interval for the wage growth? Interpret the interval.*
* *Why is the prediction interval wider than the confidence interval?*

Caution sign icon Answer the questions in a paragraph response. Remove all questions and this note before submitting! Do not include R code in your report.

## **6. Conclusion**

Based on the analysis we have done, all three of the models are solid and worth using. They all will give us a good indication of what wage growth will do based on the different variables involved. It is interesting that our estimated value based on the assumed inputs were so close on each of the 3 models performed even though we just used inflation in the first model, inflation and GDP in the second and inflation and the status of the economy in the third. We can see that the prediction intervals and confidence intervals were also very similar. Our analyses show us that different variables do affect others, and that by understanding a change in one variable of the economy we can predict to a high probability how it will affect another.