# MAT 303 Module Three Problem Set Report

Second Order Models

Bailey Davidson

Bailey.davidson@snhu.edu

Southern New Hampshire University

## **1. Introduction**

*Discuss the statement of the problem with regard to the statistical analyses that are being performed. Address the following questions in your analysis:*

* *What is the data set that you are exploring?*
* *How might your results be used?*
* *What type of analyses will you be running in this problem set?*

## **2. Data Preparation**

*There are some important variables that you have been asked to analyze in this problem set. Identify and explain these variables. Address the following questions in your analysis:*

* *What are the important variables in this data set?*
* *How many rows and columns are present in this data set?*

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

### Correlation Analysis

* *Create a scatterplot of* **wage growth and unemployment***. Comment on the relationship observed from the graph. Do you think a first order or a second order model is appropriate in this case? Explain your reasoning.*

A screen shot of a graph

AI-generated content may be incorrect.

### Reporting Results

* *Write the general form* *and the prediction equation of the second-order regression model for* **wage growth** *using* **unemployment** *as the predictor variable.*
* *Create this second-order regression model for wage growth using unemployment as the independent variable. 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.*
* *Interpret the beta estimates for the terms and 2 (squared).*

General Form:

Prediction Equation:

Where y is the predicted value of wage growth and x is unemployment, are estimates of , respectively.

Prediction equation with the outputs from my R code the new prediction model is:

The values for R-squared and adjusted R-squared are 0.9436 and 0.9424, respectively. This tells us the proportion of the variance in wage growth that is explained by the regression model. So, we are simply looking at the percentage that is predictable by the model (regression).

### 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.*

Null Hypothesis: The change in horsepower, quarter mile time, rear axle ratio, and/or whether the car is automatic vs. manual does not significantly affect fuel economy.

Alternative Hypothesis: The change in horsepower, quarter mile time, rear axle ratio, and/or whether the car is automatic vs. manual significantly affects fuel economy.

P-Value: If the p-value is less than 0.05 (5%) we reject the null hypothesis and conclude the model is significant.

Conclusion: This model is significant at a 5% level of significance, with the t-value for the model being less than 2.2e-16. Each coefficient is significant at 0.1%, with both the *intercept* and *unemployment* having t-values of less than 2e-16, and *unemployment-squared* is 6.07e-15.

### Making Predictions Using Model

* *What is the predicted wage growth if unemployment is 2.54?*
* *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.*

The prediction interval for wage growth is [7.1249, 9.358] and the confidence interval is [8.1177, 8.3651]. The prediction interval says that I am 95% confident that when the unemployment rate is 2.54, the wage growth will be between 7.1249 and 9.358. On the other hand, the confidence interval says that I am 95% confidence that average, mean, wage growth will be between 8.1177 and 8.3651 when the unemployment rate is 2.54.

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

### Reporting Results

* *Write the general form and the prediction equation of the complete second order regression model for* **wage growth** *as the* response *variable, and* **unemployment and GDP** *growth as* predictor variables*.*
* *Create this second order regression model for wage growth as the response variable, and unemployment and GDP growth as predictor variables. 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.*
* *Interpret the beta estimates for GDP2 (GDP squared) and unemployment 2 (unemployment squared).*

General Form:

Prediction Equation:

Where y is the predicted value of wage growth and is unemployment, and is GDP growth.

, , , are estimates of , , respectively.

Prediction equation with the outputs from my R code the new prediction model is:

The values for R-squared and adjusted R-squared are 0.9587 and 0.9565, respectively. This tells us the proportion of the variance in wage growth that is explained by the regression model. So, we are simply looking at the percentage that is predictable by the model (regression).

The beta estimate for GDP squared is -0.006599 with a p-value of 0.12815. This means that…however it is not significant. On the other hand, unemployment squared is 0.037685 with a p-value of 0.00489, making it significant at 1%. This means that…

### Evaluating Model Significance

* *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.*

Null Hypothesis:

Alternative Hypothesis:

P-Value: If the p-value is less than 0.05 (5%) we reject the null hypothesis and conclude the model is significant.

Conclusion: This model is significant at a 5% level of significance with the t-value being at less than 2.2e-16. Each coefficient is significant, with both the *intercept* and *unemployment* having p-values of 5.30e-14 and 8.26e-06, making them both significant at 0.1%. Also, *unemployment-squared* is 0.00489 and *gdp* is 0.04682 making them significant at 1% and 5%, respectively. The terms that are not significant at 5% based on these t-values are *gdp-squared* and *unemployment:gdp*.

### Making Predictions Using Model

* *What is the predicted wage growth if unemployment is 2.50 and GDP growth is 6.50?*
* *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.*

The predicted wage growth if unemployment is 2.50 and GDP growth is 6.50is 7.806 units.

The prediction interval for wage growth is [6.8234, 8.7886] and the confidence interval is [7.6195, 7.9925]. The prediction interval says that I am 95% confidence that when the unemployment rate is 2.50 and the GDP growth is 6.50, the wage growth will be between 6.8234 and 8.7886. On the other hand, the confidence interval says that I am 95% confidence that average, mean, wage growth will be between 7.6195 and 7.9925 when the unemployment and GDP are at these rates.

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

### Reporting Results

* *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.*

General Form:

Prediction Equation:

Where y is the predicted value of wage growth and is unemployment, and is GDP growth.

, , , are estimates of , , respectively.

Prediction equation with the outputs from my R code the new prediction model is:

The values for R-squared and adjusted R-squared are 0.9475 and 0.9446, respectively. This tells us the proportion of the variance in wage growth that is explained by the regression model. So, we are simply looking at the percentage that is predictable by the model (regression).

### Evaluating Model Significance

* *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.*

Null Hypothesis:

Alternative Hypothesis:

P-Value: If the p-value is less than 0.05 (5%) we reject the null hypothesis and conclude the model is significant.

Conclusion: This model is significant at a 5% level of significance with the t-value being at less than 2.2e-16. Each coefficient is significant, with both the *intercept* and *unemployment* having p-values of less than 2e-16, and *unemployment*-*squared* has a value of 1.24e-06, making them significant at 0.1%. Also, *unemployment:economyrecession* and *economyrecession* has 0.0272 and 0.0142 making them significant at 5%,. Lastly, *economyrecession*:*economyrecession-squared* is 0.0512 making it significant at 10%, but not 5%. This is the only term not significant at 5% based on these t-values.

### Making Predictions Using Model

* *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?*

The predicted wage growth if unemployment is 2.50 and economy is not in recession is 8.3132 units.

The prediction interval for wage growth is [7.2171, 9.4094] and the confidence interval is [8.1827, 8.4437]. The prediction interval says that I am 95% confidence that when the unemployment rate is 2.50 and the GDP growth is 6.50, the wage growth will be between 7.2171 and 9.4094. On the other hand, the confidence interval says that I am 95% confidence that average wage growth will be between 8.1827 and 8.4437 when unemployment and GDP are at these rates.

Prediction intervals are wider than confidence intervals, such as in our example, because they include an additional source of uncertainty, the random variation of individual observations, in this case individual points of measurement for the terms. Rather than just including the uncertainty in estimating the mean, which is the confidence intervals. (Forthofer et al., 2007)

## **6. Conclusion**

*Describe the results of the statistical analyses and address the following questions:*

* *Based on the analysis that you have performed and assuming that the sample size is sufficiently large, would you recommend using this model? Why or why not?*
* *Fully describe what these results mean in your scenario using proper statistical terms and concepts.*
* *What is the practical importance of the analyses that were performed?*

## **7. Citations**

Author's Last Name, First Initial. Middle Initial. (Year of Publication). Title of book: Subtitle of book, edition. Place of Publication: Publisher.

Forthofer, R. N., Lee, E. S., & Hernandez, M. (2007). Interval estimation. In *Biostatistics* (pp. 169–212). doi:10.1016/b978-0-12-369492-8.50012-1